

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 23, 2005, 18:59:31 ; Search time 169 Seconds
(Without alignments)

US-10-087-190-3
Perfect score: 1047
Sequence: 1 NSKKKGISAEEKKRTTRMMEIF.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 38670381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 16Dec04;*
1: geneseq1980s;*
2: geneseqp1990s;*
3: geneseqp2000s;*
4: geneseqp2001;*
5: geneseqp2002s;*
6: geneseqp2003as;*
7: geneseqp2003bs;*
8: geneseqp2004s;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1047	100.0	205	7 ADD84537
2	1047	100.0	205	7 ADD70015
3	1047	100.0	205	8 ADM83851
4	1047	100.0	205	8 ADM83810
5	1047	100.0	205	8 ADM83812
6	1047	100.0	205	8 ADM83793
7	1047	100.0	205	8 ADM83811
8	1047	100.0	205	8 ADM83804
9	1047	100.0	219	6 ABP75541
10	1036.5	99.0	206	8 ADM83834
11	1036.5	99.0	206	8 ADM83835
12	1011	96.6	198	8 ADM83814
13	975	93.1	190	4 ADD40043
14	975	93.1	190	7 ADD84547
15	975	93.1	190	7 ADM83557
16	975	93.1	190	8 ADM83859
17	975	93.1	190	8 ADM8303
18	975	93.1	190	8 ADM83809
19	975	93.1	190	8 ADM83858
20	948	90.5	205	8 ADM83813
21	945.5	90.3	190	8 ADD84545
22	945.5	90.3	190	8 ADM83808
23	945.5	90.3	190	8 ADM8301
24	945.5	90.3	190	8 ADM83853
25	945.5	90.3	190	8 ADM83852

Adm83853 Human can

Adm83843 Human can

Adm41829 Human pol

Ad84543 121PIFL v

Ad84543 Human can

Adm83799 Human can

Adm8345 Human can

Adm83807 Human can

Adm8344 Human can

Ad84541 121PIFL v

Adm83797 Human can

Adm83806 Human can

Adm83837 Human can

Ad84539 121PIFL v

Adm83836 Human can

Adm83795 Human can

Adm83805 Human can

Adm83815 Yeast hyp

Abg12341 Novel hum

Aag74669 Human col

Adh32846 Yeast sno

Adp26646 Mouse dyn

ALIGNMENTS

RESULT 1
ADD84337
ID ADD84537 standard; protein; 205 AA.
XX
AC ADD84337;
XX
DT 29-JAN-2004 (first entry)
DE 121PIFL protein.
XX
KW 121PIFL; 121PIFL modulation; human; chromosome 4q; cytostatic; gene therapy; vaccine; cancer; immune response; immunisation.
XX
OS Homo sapiens.
XX
PN WO200295009-A2.
XX
PD 28-NOV-2002.
XX
PP 28-FEB-2002; 2002WO-US006242.
XX
PR 05-MAR-2001; 2001US-00799250.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Challita-Eid PM, Hubert RS, Raitano AB, Parisi M, Afar DEH, Ge W; Jakobovits A;
XX
DR WPI; 2003-156757/15.
N-PSDB; ADD84536.
XX
Compositon comprising a substance that modulates the status of 121PIFL, useful in diagnosing, preventing, prognosticating or treating patients with cancer that expresses 121PIFL, such as breast, colon, ovarian or lung cancer.
XX
Claim 19; Fig 2A; 285pp; English.
XX
The present invention describes a composition (I) comprising a substance that modulates the status of 121PIFL (gene and encoded protein), or a molecule that is modulated by 121PIFL, where the status of a cell that expresses 121PIFL is modulated. The human 121PIFL gene maps to chromosome 4q. (I) has cytostatic activity, and can be used in gene therapy, and in vaccines. The composition (I) can be used for diagnosing, preventing, prognosticating or treating patients with cancer that expresses 121PIFL, such as breast, colon, ovarian or lung cancer. The 121PIFL gene or its fragment can be used to elicit a humoral or cellular immune response. 121PIFL antibodies can be used in active or passive immunisation. 121PIFL

```

Query Match          100.0%;  Score 1047;  DB 7;  Length 205;
Best Local Similarity 100.0%;  Pred. No. 1.5e-88;  Mismatches 0;
Matches 205;  Conservative 0;  Indels 0;  Gaps 0;
Y
1 MSKKKGISAEKRTRMNEIFSETKOVFQLKDELEKAPKGITAMSVEVLSVLDGMV
  |||||||SABEKRTRMNEIFSETKOVFQLKDELEKAPKGITAMSVEVLSVLDGMV 60
b
1 MSKKKGISABEKRTRMNEIFSETKOVFQLKDELEKAPKGITAMSVEVLSVLDGMV
  |||||||SABEKRTRMNEIFSETKOVFQLKDELEKAPKGITAMSVEVLSVLDGMV 60
b
61 DCERIGTGSNTWYAFPSKALHARKHKLTELESQLESGSOKHSLQSKAKIGRCTEER
  |||||||DCERIGTGSNTWYAFPSKALHARKHKLTELESQLESGSOKHSLQSKAKIGRCTEER 120
b
61 DCERIGTGSNTWYAFPSKALHARKHKLTELESQLESGSOKHSLQSKAKIGRCTEER
  |||||||DCERIGTGSNTWYAFPSKALHARKHKLTELESQLESGSOKHSLQSKAKIGRCTEER 120

```

Db	141	TRIAKELSSLADOREBOLAKAVEKVKYKDCPQDVVEETRQANKVAKAAMRWTMIAFNSWA	120
Db	121	TRIAKELSSLADOREBOLAKAVEKVKYKDCPQDVVEETRQANKVAKAAMRWTMIAFNSWA	180
Qv	181	KRKGFGEENKIDRTRGIPEDFYID	180
Db	181	KRKGFGEENKIDRTRGIPEDFYID	205
	205		

RESULT 2
ID ADJ70015
XX ADJ70015 standard; protein: 205 AA
AC ADJ70015;
XX DT 06-MAY-2004 (first entry)
XXX

XX Human heat mitochondrial F
KW mitochondrial; human; screen
Huntington's disease; Osteo-
arthritis

X **leber's hereditary optic neuropathy;** **mitochondrial encephalopathy; LHON;** **myoclonic epilepsy;** **lactic acidosis and stroke;** **MELAS;** **neuroprotective; raged red fibre syndrome; MERRF; cancer;** **osteoopathic; notopic; antidiabetic; anticancer; anticonvulsant; antiarthritic.**
Homo sapiens.

WO2003087768-A2.
23-OCT-2003.

04-APR-2003; 2003WO-US010870
12-APR-2002; 2002US-0372843P
17-JUN-2002; 2002US-0389848P
20-SEP-2002; 2002US-0412418P

(MITO-) MITOKOR.
(BUCK-) BUCK INST AGE RES.
Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM, Warnock DE;

WPI; 2003-845369/78.

Identifying a mitochondrial target for drug screening assays and for treating diseases associated with altered mitochondrial function, comprising detecting a modified polypeptide in a sample and correlating with the disease.

PT Challita-Eid PM, Hubert RS, Raitano AB, Paris M, Afar DEH, Ge W;
 PI Jakobovits A;
 XX
 XX
 DR WPI; 2004-060522/06.

PT New composition comprising a substance that modulates the status of 121PIFL or a molecule that is modulated by 121PIFL, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.

PS Example 5; SEQ ID NO 61; 21LPP; English.

CC The invention relates to a composition comprising a substance that modulates the status of 121PIFL (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121PIFL where status of the cell that expresses 121PIFL is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIFL-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIFL, inhibiting growth of cancer cells that expresses 121PIFL (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIFL, generating a mammalian immune response directed to 121PIFL, inducing an immune response, monitoring 121PIFL gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of a cancer in an individual and an assay for detecting the presence of a 121PIFL-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIFL-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIFL (ADM83793). The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptide being HLA (human leukocyte antigen)-binding epitopes from 121PIFL or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIFL. The gene for 121PIFL is located on chromosome 4q. The present sequence is a 121PIFL protein (full-length or fragment).

CC Sequence 205 AA;

CC Query Match 100.0%; Score 1047; DB 8; Length 205;
 CC Best Local Similarity 100.0%; Pred. No. 1.5e-88;
 CC Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC Qy 1 MSKKKGLSAECKTRMMEFERTKDFQKOLEKIPKEKGITAMSKEVTLQSLVGDGV 60
 CC 1 HSKKKGLSAECKTRMMEFERTKDFQKOLEKIPKEKGITAMSKEVTLQSLVGDGV 60

CC Qy 61 DCERGTSNYTAFPSKALHARKHKLVELESQSESGQKHSLSQKSTEAKIGRCETEE 120
 CC 61 DCERGTSNYTAFPSKALHARKHKLVELESQSESGQKHSLSQKSTEAKIGRCETEE 120

CC Db 121 TRLAKEISSLSDOREQKAEVKYKUCDPOVVEIROANKVAKAANRWTDNIAFKSWA 180
 CC 121 TRLAKEISSLSDOREQKAEVKYKUCDPOVVEIROANKVAKAANRWTDNIAFKSWA 180

CC Qy 181 KRKFGFEEENKIDRTFGIPEDFDYID 205
 CC 181 KRKFGFEEENKIDRTFGIPEDFDYID 205

CC Db 181 KRKFGFEEENKIDRTFGIPEDFDYID 205

RESULT 4 ADM83810

TD ADM83810 standard; protein; 205 AA.
 XX XX
 AC ADM83810;
 XX XX
 DT 03-JUN-2004 (first entry)
 XX XX
 DR Human cancer gene 121PIFL protein #3.
 XX XX
 PR Human; cancer gene 121PIFL; cytostatic; cancer; chromosome 4q; HLA; human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer; colon cancer; lung cancer; pancreatic cancer; breast cancer; cervical cancer; stomach cancer; gene therapy; vaccine.
 XX XX
 OS Homo sapiens.
 XX XX
 PN US2003223997-A1.
 XX XX
 PD 04-DEC-2003.
 XX XX
 PR 28-FEB-2002; 2002US-00087190.
 XX XX
 PR 08-FEB-2001; 2001US-00773250.
 XX XX
 PA (CHAL/) CHALLITA-EID P.M.
 PA (HUBEI/) HUBERT R.S.
 PA (RAITANO/) RAITANO A.B.
 PA (FARTI/) FARTI M.
 PA (AFAR/) AFAR D.E.H.
 PA (GEWW/) GE W.
 PA (JAKO/) JAKOBOVITS A.
 XX XX
 PR Challita-Eid PM, Hubert RS, Raitano AB, Paris M, Afar DEH, Ge W;
 PI Jakobovits A;
 XX XX
 DR WPI; 2004-060522/06.
 XX XX
 PT New composition comprising a substance that modulates the status of 121PIFL or a molecule that is modulated by 121PIFL, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.
 PT Disclosure; SEQ ID NO 20; 21LPP; English.
 XX XX
 The invention relates to a composition comprising a substance that modulates the status of 121PIFL (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121PIFL where status of the cell that expresses 121PIFL is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody or a diagnostic agent to a cell that expresses 121PIFL, inhibiting growth of cancer cells that expresses 121PIFL (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIFL, generating a mammalian immune response directed to 121PIFL, inducing an immune response, monitoring 121PIFL gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIFL-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIFL-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIFL (ADM83793). The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptide being HLA (human leukocyte antigen)-binding epitope from 121PIFL or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney

CC cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
 CC cervical cancer or stomach cancer. The composition can also be used as a
 CC vaccine to treat or prevent cancer. The composition can also be used as a
 CC 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present
 XX sequence is a 121PIF1 protein (full-length or fragment). The present
 SQ Sequence 205 AA;

Query Match Best Local Similarity 100.0%; Score 1047; DB 8; Length 205;
 Matches 205; Conservative 100.0%; Pred. No. 1.5e-88; 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MSKKGKLSAERKTRMMEIFSETKVQFDLKIAPKEKGITAMSVKTEVQSLVDDGMV 60
 Qy 61 DCERIGTSNYYWAFPSKAHLHARKHKLVEVLSQSEGSKORHASLOKSIEKAKIGCTEEER 120
 Db 61 DCERIGTSNYYWAFPSKAHLHARKHKLVEVLSQSEGSKORHASLOKSIEKAKIGCTEEER 120
 Qy 121 TRIAKELSSLRQDORQKAEVKYKCDPQVVEIROANKVAKEAANRWTNDNIAKSWA 180
 Db 181 KRKGFEENKIDRTRGIPEDFDYID 205

RESULT 5
 ADM83812

XX ID ADM83812 standard; protein; 205 AA.
 AC ADM83812;

XX DT 03-JUN-2004 (first entry)
 DE Human cancer gene 121PIF1 protein #5.

XX Human; cancer gene 121PIF1; cytostatic; cancer; chromosome 4q; HLA; human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer; cervical cancer; lung cancer; pancreatic cancer; breast cancer; kidney cancer; stomach cancer; gene therapy; vaccine; Homo sapiens.
 PN US2003223997-A1.
 XX 04-DESC-2003.

XX PF 28-FEB-2002; 2002US-00087190.
 PR 08-FEB-2001; 2001US-00779250.

XX PA (CHAL,) CHALLITA-EID P M.
 PA (HUBE,) HUBERT R S.
 PA (RAIT,) RAITANO A B.
 PA (FARI,) FARIS M.
 PA (AFAR,) AFAR D E H.
 PA (GEW,) GE W DE H.
 PA (JAKO,) JAKBOVITS A.
 PI Jakobovits A;
 PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
 DR WPI: 2004-06522/06.

XX PT New composition comprising a substance that modulates the status of 121PIF1 or a molecule that is modulated by 121PIF1, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or disclosure; SEQ ID NO 22; 211PP; English.

XX XQ The invention relates to a composition comprising a substance that modulates the status of 121PIF1 (a protein encoded by a cancer expressed gene) or a molecule that is modulated by a cancer expressed 121PIF1 where status of the cell composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1 (comprising administering a polynucleotide that encodes a single 121PIF1 (comprising administering a polynucleotide that encodes a single a patient who bears cancer cells that express 121PIF1, generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a patient who has or who is suspected of having cancer, monitoring a presence of cancer in an individual and/or a diagnostic sample from a sample from a patient who has or who is suspected of having cancer, monitoring the related protein coding sequence provided that the comprises a 121PIF1-peptide given in 16 Tables (given in the specification), the peptide being HLA (human leukocyte antigen-binding epitope from 121PIF1 (ADM83793). The peptide also comprises a polynucleotide that encodes at least one splice variant. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, cervical cancer, lung cancer, pancreatic cancer, breast cancer, kidney cancer, stomach cancer, gene therapy, vaccine, to treat or prevent cancer. The composition can also be used as a sequence is a 121PIF1 protein (full-length or fragment). The present sequence 205 AA;

XX

Query Match Best Local Similarity 100.0%; Score 1047; DB 8; Length 205;
 Matches 205; Conservative 100.0%; Pred. No. 1.5e-88; 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MSKKGKLSAERKTRMMEIFSETKVQFDLKIAPKEKGITAMSVKTEVQSLVDDGMV 60
 Qy 61 DCERIGTSNYYWAFPSKAHLHARKHKLVEVLSQSEGSKORHASLOKSIEKAKIGCTEEER 120
 Db 121 TRIAKELSSLRQDORQKAEVKYKCDPQVVEIROANKVAKEAANRWTNDNIAKSWA 180
 Qy 181 KRKGFEENKIDRTRGIPEDFDYID 205
 Db 181 KRKGFEENKIDRTRGIPEDFDYID 205

RESULT 6

ADM83793

XX ID ADM83793 standard; protein; 205 AA.

XX DT 03-JUN-2004 (first entry)
 DE Human cancer gene 121PIF1 protein #1.

XX KW Human; cancer gene 121PIF1; cytostatic; cancer; chromosome 4q; HLA; human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer; colon cancer; lung cancer; pancreatic cancer; breast cancer, cancer,

KW cervical cancer; stomach cancer; gene therapy; vaccine.
 XX OS Homo sapiens.
 XX PN US2003223997-A1.
 XX PD 04-DEC-2003.
 XX PP 28-FEB-2002; 2002US-00087190.
 XX PR 08-FEB-2001; 2001US-00779250.
 XX PA (CHAL/) CHALLITA-EID P M.
 XX PA (HUBE/) HUBERT R S.
 XX PA (RAIT/) RAITANO A B.
 XX PA (FARI/) FARIS M.
 XX PA (AFAR/) AFAR D E H.
 XX PA (GEWW/) GE W.
 XX PA (JAKO/) JAKOBOWITS A.
 XX PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
 XX PI Jakobovits A;
 XX DR WPI; 2004-060522/06.
 XX DR N-PSB; ADM83792.
 XX DT N-PSB; ADM83792.
 PT New composition comprising a substance that modulates the status of 121PIF1 or a molecule that is modulated by 121PIF1, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.
 PT XX
 PS Claim 19; SEQ ID NO 3; 211pp; English.
 CC The invention relates to a composition comprising a substance that modulates the status of 121PIF1 (a protein encoded by a cancer expressed gene), or a molecule that is modulated by 121PIF1 where status of the cell that expresses 121PIF1 is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1, inhibiting growth of cancer cells that expresses 121PIF1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIF1, generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIF1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIF1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIF1 (ADM83793. The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIF1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present sequence is a 121PIF1 protein (full-length or fragment).
 CC Sequence 205 AA;

Query Match 100.0%; Score 1047; DB 8; Length 205;
 Best Local Similarity 100.0%; Pred. No. 1.5e-88;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKKGGISAEERTRTRMMEISBTKDVFQQLKOLEKAKPEKEITAMSVKEVQSLVDDGMV 60
 DB 1 MSKKGGISAEERTRTRMMEISBTKDVFQQLKOLEKAKPEKEITAMSVKEVQSLVDDGMV 60
 QY 61 DCERGTSNYWAFPSKALHARKHKLVLVQSLSGSKQHASLQSKERAKIGRCSTEER 120
 DB 61 DCERGTSNYWAFPSKALHARKHKLVLVQSLSGSKQHASLQSKERAKIGRCSTEER 120
 QY 121 TRLAKESSLADOREQKAEVYKQCDPQVVEERQANKVAKEARWNTNIAIKSWA 180
 DB 121 TRLAKESSLADOREQKAEVYKQCDPQVVEERQANKVAKEARWNTNIAIKSWA 180
 QY 181 KRKFGFEEPKIDRTFGIPEDFDYID 205
 DB 181 KRKFGFEEPKIDRTFGIPEDFDYID 205
 RESULT 7
 ID ADM83011
 ID ADM83811 standard; protein; 205 AA.
 XX AC ADM83011;
 XX AC ADM83011;
 XX DT 03-JUN-2004 (first entry)
 XX DE Human cancer gene 121PIF1 protein #4.
 XX KW Human; cancer gene 121PIF1; cytostatic; cancer; chromosome 4q; HLA; human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer; colon cancer; lung cancer; pancreatic cancer; breast cancer; cervical cancer; stomach cancer; gene therapy; vaccine.
 XX OS Homo sapiens.
 XX DN US2003223997-A1.
 XX DR 04-DEC-2003.
 XX PR 28-FEB-2002; 2002US-00087190.
 XX PR 08-FEB-2001; 2001US-00779250.
 XX PA (CHAL/) CHALLITA-EID P M.
 XX PA (HUBE/) HUBERT R S.
 XX PA (RAIT/) RAITANO A B.
 XX PA (FARI/) FARIS M.
 XX PA (AFAR/) AFAR D E H.
 XX PA (GEWW/) GE W.
 XX PA (JAKO/) JAKOBOWITS A.
 XX PI Challita-Eid PM, Hubert RS, Raitano AB, Faris M, Afar DEH, Ge W;
 XX PI Jakobovits A;
 XX DR WPI; 2004-060522/06.
 XX PT New composition comprising a substance that modulates the status of 121PIF1 or a molecule that is modulated by 121PIF1, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.
 XX Disclosure; SEQ ID NO 21; 211pp; English.
 CC The invention relates to a composition comprising a substance that modulates the status of 121PIF1 (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121PIF1 where status of the cell that expresses 121PIF1 is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal antibody, a non-human transgenic animal that produces an antibody, a hybridoma that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising

Sequence 205 AA; Long OR fragment). - Present

RESULT 8
ADM83804
ID ADM83804 **standard; protein; 205 AA.**
XX
AC ADM83804;
XX
DT 03-TUN-2004 (first entry)
XX
DE Human cancer gene 121PIF1 protein #2.
XX
KW Human; cancer gene 121PIF1; cytostatic; cancer; chromosome 4c
KW human leukocyte antigen; prostate cancer
KW colon cancer

KW cervical cancer; lung cancer; pancreatic cancer; bladder cancer; kidney cancer; cervical cancer; stomach cancer; breast cancer; Homo sapiens. JS gene therapy; vaccine. US2003223997-A1. 04-DEC-2003. 28-FEB-2002; 2002US-00087190.

THEORY AND PRACTICE IN THE FIELD OF INFORMATION TECHNOLOGY

Match	100	0%	Score 1047;	DB 8;	Length 205;
Local Similarity	100.0%	Pred. No. 1.5e-88;	Mismatches 0;	Indels 0;	Gaps 0;
1	MSKKKGSAEERTRMEISETKDVFOLKDKIETKIAKEKEGTAMSVEKYLQSLVDDGAV	MSKKKGSAEERTRMEISETKDVFOLKDKIETKIAKEKEGTAMSVEKYLQSLVDDGAV	MSKKKGSAEERTRMEISETKDVFOLKDKIETKIAKEKEGTAMSVEKYLQSLVDDGAV	MSKKKGSAEERTRMEISETKDVFOLKDKIETKIAKEKEGTAMSVEKYLQSLVDDGAV	MSKKKGSAEERTRMEISETKDVFOLKDKIETKIAKEKEGTAMSVEKYLQSLVDDGAV
61	DOERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DOERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DOERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DOERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DOERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER
61	DCERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DCERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DCERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DCERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER	DCERIGTSNYWAPSKALHARHGKUVELEQOLSGSGSOKHASTLOKSIEGAKIGRCRTER
121	TRLAKELSSLRDOREQKAEVKYKODDPQWEEIRQANKVAKAANRWTNIAFKTWA	TRLAKELSSLRDOREQKAEVKYKODDPQWEEIRQANKVAKAANRWTNIAFKTWA	TRLAKELSSLRDOREQKAEVKYKODDPQWEEIRQANKVAKAANRWTNIAFKTWA	TRLAKELSSLRDOREQKAEVKYKODDPQWEEIRQANKVAKAANRWTNIAFKTWA	TRLAKELSSLRDOREQKAEVKYKODDPQWEEIRQANKVAKAANRWTNIAFKTWA
180					

Db 121 TRLAKEISSLRQDQEQLAEVYKQCDPQVVEIROANKVAKEAANRWTNDNIAIKSWA 180
 Qy 181 KRKFGFEEENKIDRTFGIPEDDYID 205
 Db 181 KRKFGFEEENKIDRTFGIPEDDYID 205

RESULT 9

ABP75541
 ABP75541 standard; protein; 219 AA.

AC ABP75541;
 XX
 DT 10-FEB-2003 (first entry)

DE Human Beccotry polypeptide SPTM SEQ ID NO 725.

KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia; asthma; Crohn's disease; neurological disorder; epilepsy; cancer; Huntington's disease; Alzheimer's disease; Cretzfeldt-Jakob disease; multiple sclerosis; Parkinson's disease; cell proliferative disorder; anti-inflammatory; immunosuppressive; neuroprotective; neurotropic; neuroleptic; anticonvulsant; cytostatic; anti-HIV; human immunodeficiency virus; secretory polynucleotide; Beccotry protein. KW Homo sapiens.

OS Homo sapiens.

PN WO200283876-A2.

XX
 PD 24-OCT-2002.

XX
 PR 27-MAR-2002; 2002WO-US009921.

XX
 PR 29-MAR-2001; 2001US-0280067P.

PR 16-MAY-2001; 2001US-0291228P.

PR 17-MAY-2001; 2001US-0291849P.

PR 19-JUN-2001; 2001US-0293428P.

PR 20-JUN-2001; 2001US-0299776P.

PR 20-JUN-2001; 2001US-0300001P.

XX
 PR (INCY-) INCYTE GENOMICS INC.

XX
 PR Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J, Chinn J, Dufour GE, Hillman JL, Yu JY, Thasos O, Yap PS, Aushey SR, Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstain EH, Peralta CH, David MR, Lewis SA, Chen AJ, Panzer SR, Harris B, Flores V, Marwaha R, Lo A, Lan RV, Uraslka ME.

XX
 PR WPI; 2003-07553/07.

DR N-PSDB; AB235987.

XX
 PT New human secretory proteins and polynucleotides, useful for diagnosing, treating, preventing or detecting autoimmune/inflammatory disorders (e.g. Alzheimer's), or cell proliferations or cancers.

PT The invention relates to a secretory polynucleotide sequence (designated sptm) comprising any of 557 polynucleotide sequences (AB235837-AB236403), a naturally occurring polynucleotide sequence at least 90 % identical to the polynucleotide sequence, a polynucleotide complementary to them or an RNA equivalent of them. The polypeptide or polynucleotide are useful for treating, preventing or diagnosing a disease or condition associated with the expression of functional SPTM. These are particularly useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia, or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythaemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). The present sequence is one of the SPTM proteins of the invention (ABP7538/ABP75962). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/published_pct_sequences

XX Sequence 219 AA;

Query Match 100 0%; Score 1047; DB 6; Length 219; Best Local Similarity 100 0%; Pred. No. 1; 7.e-80; Indels 0; Gaps 0; Matches 205; Conservative 0; Mismatches 0;

Qy 1 MSKKKGISAEERTRMMBISFTKDVQLKOLEKIAPKKEKITAMSVKEVQSLVQDGMV 60
 Db 15 MSKKKGISAEERTRMMBISFTKDVQLKOLEKIAPKKEKITAMSVKEVQSLVQDGMV 74

Qy 61 DCERIGHSNYWAFPSKALHARKHKLVELESQSEGSKQHSLQSKIEAKIGRCETEE 120
 Db 75 DCERIGHSNYWAFPSKALHARKHKLVELESQSEGSKQHSLQSKIEAKIGRCETEE 134

Qy 121 TRLAKEISSLRQDQEQLAEVYKQCDPQVVEIROANKVAKEAANRWTNDNIAIKSWA 180
 Db 135 TRLAKEISSLRQDQEQLAEVYKQCDPQVVEIROANKVAKEAANRWTNDNIAIKSWA 194

Qy 181 KRKFGFEEENKIDRTFGIPEDDYID 205
 Db 195 KRKFGFEEENKIDRTFGIPEDDYID 219

RESULT 10

ADMB83334 standard; protein; 206 AA.

XX
 PR 03-JUN-2004 (first entry)

DE Human cancer gene 121PIF1 variant protein #1.

XX
 KW Human; cancer gene 121PIF1; cytostatic; cancer; chromosome 4q; HLA; human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;

KW colon cancer; lung cancer; pancreatic cancer; breast cancer; cervical cancer; stomach cancer; gene therapy; vaccine.

OS Homo sapiens.

XX
 PR US2003223997-A1.

XX
 PD 04-DEC-2003.

XX
 PR 28-FEB-2002; 2002US-00087190.

XX
 PR 08-FEB-2001; 2001US-00779250.

XX
 PR (CHALI/) CHALLITA-BID P M.

PA (HUBER/) HUBERT R S.

PA (RAIT/) RAITANO A B.

PA (FARI/) FARIS M.

PA (AFAR/) AFAR D E H.

PA (GEW/) GE W.

PA (JAKO/) JAKOBOWITS A.

XX
 PR Challita-Bid PM, Hubert RS, Raitano AB, Faris M, Afar DBH, Ge W;

PR Jakobowitz A;

XX
 PR WPI; 2004-060522/06.

XX
 PT New composition comprising a substance that modulates the status of 121PIF1 or a molecule that is modulated by 121PIF1, useful for detecting,

Query Match 99.0%; Score 1036.5; DB 8; Length 206;
 Best Local Similarity 99.5%; Pred. No. 1.4e-87; Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 NSKKKGLSAAEKRTRMMEIFSETKDVQFOLKLEKIAPEKEKITAMSVKEVIVQSLVNDGMV 60
 1 NSKKKGLSAAEKRTRMMEIFSETKDVQFOLKLEKIAPEKEKITAMSVKEVIVQSLVNDGMV 60

CC antibody, a non-human transgenic animal that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1, inhibiting growth of cancer cells that expresses 121PIF1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIF1. Generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIF1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIF1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIF1 (ADM8793. The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIF1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present sequence is a 121PIF1 protein (full-length or fragment).

CC antibody, a non-human transgenic animal that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1, inhibiting growth of cancer cells that expresses 121PIF1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIF1. Generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIF1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIF1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIF1 (ADM8793. The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIF1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present sequence is a 121PIF1 protein (full-length or fragment).

DB RESULT 12
 ADM83814 standard; protein; 198 AA.
 XX
 AC ADM83814;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Human cancer gene 121PIF1 protein fragment.
 XX
 KW Human; cancer gene 121PIF1; cytostatic; cancer; chromosome 4q; HLA; human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer; colon cancer; lung cancer; pancreatic cancer; breast cancer; cervical cancer; stomach cancer; gene therapy; vaccine.
 KW Homo sapiens.
 OS US2003223997-A1.
 XX
 PD 04-DEC-2003.
 XX
 PP 28-FEB-2002; 2002US-00087190.
 XX
 PR 08-FEB-2001; 2001US-00779250.
 XX
 PA (CHALI/) CHALLITA-EID P M.
 PA (HUBE/) HUBERT R S.
 PA (RAIT/) RAITANO A B.
 PA (FARI/) PARIS M.
 PA (AFAR/) AFAR D E H.
 PA (GEW/) GE W.
 PA (JAKO/) JAKOBOWITS A.
 XX
 PI Challita-Bid PM, Hubert RS, Raitano AB, Paris M, Afar DEH, Ge W;
 PI Jakobowits A;
 XX
 DR WPI; 2004-060522/06.

PT New composition comprising a substance that modulates the status of 121PIF1 or a molecule that is modulated by 121PIF1, useful for detecting, treating or preventing cancer e.g. prostate, bladder, colon, breast or lung cancer.

PT Example 2; SEQ ID NO 24; 211P; English.

PT The invention relates to a composition comprising a substance that modulates the status of 121PIF1 (a protein encoded by a cancer expressed gene) or a molecule that is modulated by 121PIF1 where status of the cell that expresses 121PIF1 is modulated. Also included are a pharmaceutical composition comprising the novel composition in a human unit dose form, a recombinant protein comprising an antigen-binding region of a monoclonal

Query Match 99.0%; Score 1036.5; DB 8; Length 198;
 Best Local Similarity 99.0%; Pred. No. 3.1e-85; Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 KGLSAAEKRTRMMEIFSETKDVQFOLKLEKIAPEKEKITAMSVKEVIVQSLVNDGMV 64
 1 KGLSAAEKRTRMMEIFSETKDVQFOLKLEKIAPEKEKITAMSVKEVIVQSLVNDGMV 60

CC antibody, a non-human transgenic animal that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1, inhibiting growth of cancer cells that expresses 121PIF1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIF1. Generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIF1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIF1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIF1 (ADM8793. The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIF1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present sequence is a 121PIF1 protein (full-length or fragment).

CC antibody, a non-human transgenic animal that produces an antibody, a single chain monoclonal antibody that immunospecifically binds to a 121PIF1-related protein (comprising the variable domains of the heavy and light chains of a monoclonal antibody), a vector comprising a polynucleotide that encodes a single chain monoclonal antibody, a polynucleotide that encodes an analogue peptide, delivering a cytotoxic agent or a diagnostic agent to a cell that expresses 121PIF1, inhibiting growth of cancer cells that expresses 121PIF1 (comprising administering to the cells the composition), treating a patient who bears cancer cells that express 121PIF1. Generating a mammalian immune response directed to 121PIF1, inducing an immune response, monitoring 121PIF1 gene products in a biological sample from a patient who has or who is suspected of having cancer, monitoring the presence of cancer in an individual and an assay for detecting the presence of a 121PIF1-related protein or polynucleotide in a biological sample from a patient who has or who is suspected of having cancer. The composition may comprise a polynucleotide that comprises a 121PIF1-related protein coding sequence provided that the coding sequence does not encode the entire amino acid sequence of 121PIF1 (ADM8793. The substance also comprises a polynucleotide that encodes at least one peptide given in 16 Tables (given in the specification), the peptides being HLA (human leukocyte antigen)-binding epitopes from 121PIF1 or its splice variants. The composition is useful for detecting, treating or preventing cancer, preferably prostate cancer, bladder cancer, kidney cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer, cervical cancer or stomach cancer. The composition can also be used as a vaccine to treat or prevent cancer that expresses or overexpresses 121PIF1. The gene for 121PIF1 is located on chromosome 4q. The present sequence is a 121PIF1 protein (full-length or fragment).

DB RESULT 13
 ID AAM40043 standard; protein; 190 AA.
 XX
 AC AAM40043;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 3188.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
 XX
 OS Homo sapiens.
 PN WO200153312-A1.

RESULT 15
 ADM83857
 ID ADM83857 standard; protein; 190 AA.
 XX
 AC ADM83857;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Human cancer gene 121P1F1 protein 16-205.
 XX
 KW Human; cancer gene 121P1F1; cytostatic; cancer; chromosome 4q; HLA;
 KW human leukocyte antigen; prostate cancer; bladder cancer; kidney cancer;
 KW colon cancer; lung cancer; pancreatic cancer; breast cancer;
 KW cervical cancer; stomach cancer; gene therapy; vaccine.
 OS Homo sapiens.
 XX
 PN US2003223997-A1.
 XX
 PD 04-DEC-2003.
 XX
 PP 28-FEB-2002; 2002US-00087190.
 XX
 PR 08-FEB-2001; 2001US-0077250.
 XX
 PA (CHAL/) CHALLITA-EID P M.
 PA (HUBE/) HUBERT R S.
 PA (RAIT/) RAITANO A B.
 PA (FARI/) PARIS M.
 PA (AFAR/) AFAR D E H.
 PA (GEW/) GE W.
 PA (JAKO/) JAKOBOWITS A.
 XX
 PI Challita-Eid PM, Hubert RS, Raitano AB, Paris M, Afar DEH, Ge W;
 PI Jakobowits A;
 XX
 DR WPI; 2004-060522/06.
 XX
 PT New composition comprising a substance that modulates the status of
 PT 121P1F1 or a molecule that is modulated by 121P1F1, useful for detecting,
 PT treating or preventing cancer e.g. prostate, bladder, colon, breast or
 PT lung cancer.
 XX
 PS Example 5; SEQ ID NO 67; 211pp; English.

CC The invention relates to a composition comprising a substance that
 CC modulates the status of 121P1F1 (a protein encoded by a cancer expressed
 CC gene) or a molecule that is modulated by 121P1F1 where status of the cell
 CC that expresses 121P1F1 is modulated. Also included are a pharmaceutical
 CC composition comprising the novel composition in a human unit dose form, a
 CC recombinant protein comprising an antigen-binding region of a monoclonal
 CC antibody, a non-human transgenic animal that produces an antibody, a
 CC hybridoma that produces an antibody, a single chain monoclonal antibody
 CC that immunospecifically binds to a 121P1F1-related protein (comprising
 CC the variable domains of the heavy and light chains of a monoclonal
 CC antibody), a vector comprising a polynucleotide that encodes a single
 CC chain monoclonal antibody, a polynucleotide that encodes an analogue
 CC peptide, delivering a cytotoxic agent or a diagnostic agent to a cell
 CC that expresses 121P1F1, inhibiting growth of cancer cells that expresses
 CC 121P1F1 (comprising administering to the cells the composition), treating
 CC a patient who bears cancer cells that express 121P1F1, generating a
 CC mammalian immune response directed to 121P1F1, inducing an immune
 CC response, monitoring 121P1F1 gene products in a biological sample from a
 CC patient who has or who is suspected of having cancer, monitoring the
 CC presence of cancer in an individual and an assay for detecting the
 CC presence of a 121P1F1-related protein or polynucleotide in a biological
 CC sample from a patient who has or who is suspected of having cancer. The
 CC composition may comprise a polynucleotide that comprises a 121P1F1-
 CC related protein coding sequence provided that the coding sequence does
 CC not encode the entire amino acid sequence of 121P1F1 (ADM83793. The
 CC substance also comprises a polynucleotide that encodes at least one
 CC peptide given in 16 Tables (given in the specification), the peptides
 CC being HLA (human leukocyte antigen)-binding epitopes from 121P1F1 or its

CC splice variants. The composition is useful for detecting, treating or
 CC preventing cancer, preferably prostate cancer, bladder cancer, kidney
 cancer, colon cancer, lung cancer, pancreatic cancer, breast cancer,
 CC cervical cancer or stomach cancer. The composition can also be used as a
 CC vaccine to treat or prevent cancer that expresses or overexpresses
 CC 121P1F1. The gene for 121P1F1 is located on chromosome 4q. The present
 CC sequence is a 121P1F1 protein (full-length or fragment).
 XX
 Sequence 190 AA:
 CC
 Query Match 93.1%; Score 975; DB 8; Length 190;
 CC Best Local Similarity 100.0%; Pred. No. 6.3e-82;
 CC Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Db 16 MMEIFSTKDVFOLKDJKEKLAPEKEGTAMSKEVITQSLVUDGMVDCERIGTSNYWWAfp 75
 CC
 Db 1 MMEIFSTKDVFOLKDJKEKLAPEKEGTAMSKEVITQSLVUDGMVDCERIGTSNYWWAfp 60
 CC
 Qy 76 SKALHAKKHKLVELESQLESGSOKHSLQKSTEAKIGRCETEERTRIAKELSSRQRE 135
 CC
 Db 61 SKALHAKKHKLVELESQLESGSOKHSLQKSTEAKIGRCETEERTRIAKELSSRQRE 120
 CC
 Qy 136 QIKAEVFKYKQDPOVNEIQRANKVAKAEEAARNWTNDIPIAKSWAKRKGFGREENKDRfp 195
 CC
 Db 121 QIKAEVFKYKQDPOVNEIQRANKVAKAEEAARNWTNDIPIAKSWAKRKGFGREENKDRfp 180
 CC
 Qy 196 GIPEDFDPYID 205
 CC
 Db 181 GIPEDFDPYID 190
 CC
 Search completed: March 23, 2005, 19:14:48
 Job time : 171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2005, 19:09:32 ; Search time 43 Seconds
(without alignments)

335.885 Million cell updates/sec

Title: US-10-087-190-3
Perfect score: 1047
Sequence: 1 MSKKKGISAEKKRTRMMBIP.....FEENKIDRTFGIPEDFDYID 205

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 7464964 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A COMB.pep: *
2: /cgn2_6/prodata/1/iaa/5B COMB.pep: *
3: /cgn2_6/prodata/1/iaa/6A COMB.pep: *
4: /cgn2_6/prodata/1/iaa/6B COMB.pep: *
5: /cgn2_6/prodata/1/iaa/PCITUS COMB.pep: *
6: /cgn2_6/prodata/1/iaa/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1
US-09-621-976-4959
; Sequence 4959, Application US/09621976
; Patent No. 65339063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jober, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054922
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4959
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-621-976-4959

Query Match 56.7%; Score 593.5; DB 4; Length 127;
Best Local Similarity 89.6%; Pred. No. 1.7e-50;
Matches 120; Conservative 5; Mismatches 2; Indels 7; Gaps 1;

Qy 1 MSKKKGISAEKKRTRMMBIP.....FEENKIDRTFGIPEDFDYID 60

Db 1 MSKKKGISAEKKRTRMMBIP.....FEENKIDRTFGIPEDFDYID 60

Qy 61 DCERIGSNYWAFFSKALHARKHKKLVEFQSLQSKGSKQHASLQSKIEKAKIGRCETEE 120
Db 61 DCERIGSNYWAFFSKALHARKHKKLVEFQSLQSKGSKQHASLQSKIEKAKIGRCETEE 117

Qy 121 TRLAKEISSLRDRQ 134

Db 118 ---IKLSGMQEE 127

RESULT 2
US-09-248-796A-20276
; Sequence 63, Appl
; Sequence 10, Appl
; Sequence 14485, Application US/09248796A
; Patent No. 10747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196-132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13

28 105 10.0 251 4 US-09-914-239-64
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

29 105 10.0 284 4 US-09-914-239-60
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

30 104 9.9 1939 4 US-09-938-092-915
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

31 104 9.9 1939 4 US-09-949-016-1104
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

32 102 9.7 281 4 US-09-914-239-45
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

33 102 9.7 284 4 US-09-914-239-47
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

34 102 9.7 284 4 US-09-914-239-57
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

35 102 9.7 372 1 US-07-813-594A 3
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

36 102 9.7 372 1 US-08-330-515-3
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

37 102 9.7 1031 4 US-09-914-239-24
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

38 102 9.7 1940 4 US-09-538-092-901
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

39 102 9.7 1963 4 US-09-949-016-8888
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

40 101.5 9.7 2662 4 US-09-634B-31
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

41 101.5 9.7 2663 4 US-09-538-092-122
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

42 101.5 9.6 284 4 US-09-914-239-59
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

43 100 9.6 1104 3 US-08-923-992A-4
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

44 100 9.6 1164 3 US-08-923-992A-10
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

45 99.5 9.5 588 4 US-08-714-711-42
Sequence 64, Appl
Sequence 60, Appl
Sequence 915, Appl
Sequence 11104, A
Sequence 45, Appl
Sequence 47, Appl
Sequence 57, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 24, Appl
Sequence 901, Appl
Sequence 8888, Appl
Sequence 31, Appl
Sequence 1252, Appl
Sequence 59, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 42, Appl

; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO: 14485
 ; LENGTH: 260
 ; TYPE: PRT
 ; ORGANISM: *Candida albicans*
 ; US-09-248-736A-14485
 Query Match
 Best Local Similarity 18.3%; Score 191.5; DB 4; Length 260;
 Matches 59; Conservative 47; MisMatches 75; Indels 37; Gaps 8;
 Qy 1 MSKKGKSLSAEKKRMMELFSETKDVFOLKOLEKAPKGKGKGTAMSVPVQIQLSVDGMV 60
 Db 60 MPPKGKLTQEEKUSALLWEQSDHMTYTKBIESKASKOCKIPPMQMKELVIALVEGLV 119
 Qy 61 DCERIGTNYWYMAPSKALHARKHKEVIEQSESGSKHASKLQSIKEKAK-----IGR 114
 Db 120 EDRGCTINLYWSP----YLOHQKQ-----QFTHDRNRTIANTEEDSLICR 165
 Qy 115 C--ETEERTRLAKERSSLR-----DORBKAKVEVKYKODCPQVVEETRQANVAKA 166
 Db 166 CKDGETGVRNQHERASKIRFCDSLERIDSQOSLOSLKOSES--VE----NLVLSLAF 218
 Qy 167 NEWTONIAFKSWAKRKGOFEEENKIDTGFPGFEDDVI 204
 Db 219 --FSDSDDICILYSLRQGLTMILKTFPSLPLEFEEI 254
 RESULT 3
 US-09-538-092-1321
 ; Sequence 1321 Application US/09538092
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; TITLE OF INVENTION: Mansfield, Traci A.
 ; FILE REFERENCE: Protein-Protein Complexes and Method of Using Same
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CurapatSeqFormatter Version 0.9
 ; SEQ ID NO: 1321
 ; LENGTH: 1270
 ; TYPE: PRT
 ; ORGANISM: *Homo sapiens*
 ; FEATURE: misc_feature
 ; NAME/KEY: misc_feature
 ; LOCATION: (0) .(0)
 ; OTHER INFORMATION: Polypeptide Accession Number_Q14203
 ; US-09-538-092-1321
 Query Match
 Best Local Similarity 12.8%; Score 134.5; DB 4; Length 1270;
 Matches 50; Conservative 43; Mismatches 106; Indels 15; Gaps 5;
 Qy 1 SKKKGKSLSAEKKRMMELFSETKDVFOLKOLEKAPKGKGKGTAMSVPVQIQLSVDGMV 61
 Db 271 ARKEAKALAKERKERYNEEMADTAIAEMATLKKMEEAERESLQEEVAKERVDSLTD 330
 Qy 62 CERI-----GTSNYWYMAPSKALHARKHKEVIEQSESGSKHASKLQSIKEA 110
 Db 331 LEIJKALEBEGKSDGMASSYQKQLEONARLKOALVNRDISSSERKQEHKLQKLMK- 389
 Qy 111 KIGRCET--EERLAKELSSRDRQEEQKAEVYKQCDPOVEREIRQANVAKA 168
 Db 390 KNGELEVVRQQRERLQELSQESTIDELKEQDAALGAE-EMVEMLTDRNLNLJEKVRE 649
 ; US-09-104-324B-4
 Query Match
 Best Local Similarity 11.2%; Score 117; DB 3; Length 976;
 Matches 63; Conservative 38; MisMatches 85; Indels 78; Gaps 13;
 Qy 3 KKKGKSLSAEKKRMMELFSETKDVFOLKOLEKAPKGKGKGTAMSVPVQIQLSVDGMV 51
 Db 413 OKKSELEEMTKLNKEVLEELKKVKGKETLYENKQKESKIALKG-TEOBIGL 471
 Qy 91 -----SOLSESGSKHASKLQSIKEA 520
 Db 472 QARKEVHDILQIOLTAITSEQYKSEVQKDLTTELNEKLNTELTSHCNKSLLENKLT 531
 Qy 52 OSI-----VDPGMVDCERIGTNYWYMAPSKALHARKHKEVIEQSESGSKHASKLQSIKEA 136
 Db 532 QBTSMTLELKNOQDINNNKQOERMLQI--NQETETORNELBYVREBKORDE 589
 Qy 137 KRAEVYKVD-CD-----PQVVEETRQANVAKA--AARWTDNIFAIK----- 177
 Db 590 VKCLDKSSEENNLRKQVVENTKVKYIEELQOENKALKKGTAESKOLNVBEKVKLE 649
 Qy 178 -SWAKRKG-----FRENKI 191
 Db 650 LSAKOKRGEITDYZQKEIEDK 673
 RESULT 5

US-00-538-092-1339
; Sequence 1339, Application US/00538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/00/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO: 1339
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number Q15431
; US-09-538-092-1339
Query Match 11.2%; Score 117; DB 4; Length 976;
Best Local Similarity 23.9%; Pred. No. 0.0099; Matches 63; Conservative 38; Mismatches 85; Indels 78; Gaps 13; Gapopen 0; Query 3 KKKGGLSAAE--KRTRMMEI-SETKOV-----FOLKDLIKEKIAKEKGITAMSVKEVL 51
Db 413 QKKSSELEEMWTKTNNKEVRELKVLGERBTLYENKOPEKABELKG-TEQELIGL 471
Qy 52 QSL--VDDGNDCERGTSNYWAFPSKAL-----HARKHKJLEVLE-- 90
Db 472 QAREKEVHDJELQIOLTAITSPHOYSKEVKDQIKTELENEKLNTEFLSHCNLUSLENKELT 531
Qy 91 -----SOLSEGSKQHASIQLQSKTEKAKIGRCETEERTRLAKELSSLRDOREQ 136
Db 532 QETSDMTELEKQNOQEDINNNKQOERMLKQI-E-NIQETEQLRNLEYRRELKOKRDE 589
Qy 137 IKADEVKYKQ-CD-----POVVEEIRGQANKVAK-E-ANRWDNIFAIK----- 177
Db 590 VKCKLUDKSEENCNLRKQVENKVKYIIELOOENKALKKGTAESKOLNVYEVKUNKLE 649
Qy 178 -SWAKRKG-----FEENKI 191
Db 650 LESAQOKKFGEITDTVQKEIEKKI 673
; RESULT 6
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/00/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 62
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-914-259-62
Query Match 10.8%; Score 113.5; DB 4; Length 284;
Best Local Similarity 21.1%; Pred. No. 0.0042; Matches 48; Conservative 48; Mismatches 78; Indels 53; Gaps 8; Gapopen 0; Query 2 SKKKGLSA-EKRTRMMEISETKOVFQLKDLIKEKIAKEKGITAMSVK-BVLSQLYDGM 59
Db 45 AKEKLRAASDERDVTLEELHKAEDSLLAAD-ETAKAEADAVASLNRIQVLEELRAQ 103
Qy 60 VDCERITGTSNYWAFPSKALHARKHKJLEVLEQSEGSK----- 99
Db 104 --ERLATALOKLKEEAKADESERGMKVIESRAQDEEKMEOIQLKKEAHEDADR 160
Qy 100 -----HASIQLSKTEAKI--GRETEERTRLAKELSSLRDOQKARVYK 145
Db 161 KYEEVARKLVIESDIEBRAEAEGLSGKC---AEFLERELKTVNLKSLDQAEKYS 215
Qy 146 DCDPQVVEETRQANKVAKEAANRWTNDNIAFKSWAKRKGFEENKID 192
Db 216 QKEDKVEEIKVLSDKLEKAETR-----AEFAERSVTKLEKSID 254
; RESULT 7
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; APPLICANT: Makowski, Lee
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/00/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 55
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Sus scrofa
; US-09-914-259-55
Query Match 10.9%; Score 114; DB 4; Length 284;
Best Local Similarity 23.2%; Pred. No. 0.0037;
; RESULT 8
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107136.132
; CURRENT APPLICATION NUMBER: US/00/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13

PRIOR FILING DATE: US 6/0/096, 409
 NUMBER OF SEQ ID NOS: 28208
 LENGTH: 630
 TYPE: PRT
 ORGANISM: Candida albicans
 US-09-248-796A-20275

Query Match Best Local Similarity 10.7%; Score 112.5; DB 4; Length 630; Matches 49; Conservative 43; Mismatches 79; Indels 35; Gaps 8; Qy 3 KKGGLSABERKRTRMMEIFSETKDVFOLKLEKAPKKGIT-AMVKEVILQSLVUDGNY 60 Db 158 KTKNSDDEBLKQKQLELEKVK---IDQTADEKKGITREALKBETVKNSG-- 210 Qy 61 DCERIGTSNYYWAFPS-KALHARHKLVEU---BSQLEBGSOKRASLQKST----- 107 Db 211 ---LSTTSELALAKTQVLSLEKEKESLQFLSGNKSKELEDLYIQHSDISEKUKALTDL 265 Qy 108 EKAKIGRCETEERTRLAKELLSRDRQLEKAEVYKODDQVWVERIRQANKVKEAA 166 Db 267 KEKIQFDFDSKQKLTELENDLSTKCELETEKTOTSKENLERKDEIVKLNKHELEK 326 Db 327 NT---DN---SGAKKELLERVKSL 344

RESULT 9
 US-09-914-259-43
 ; Sequence 43, Application US/09914259
 ; Patent No. 6495336
 ; GENERAL INFORMATION:
 ; APPLICANT: Makowski, Lee
 ; APPLICANT: Hyman, Paul
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
 ; FILE REFERENCE: 8471-010-939
 ; CURRENT APPLICATION NUMBER: US/09/914,259
 ; NUMBER OF SEQ ID NOS: 180
 ; SEQ ID NO: 43
 ; LENGTH: 284
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-914-259-43

Query Match Best Local Similarity 10.7%; Score 112.5; DB 4; Length 630; Matches 49; Conservative 43; Mismatches 79; Indels 35; Gaps 8; Qy 3 KKGGLSABERKRTRMMEIFSETKDVFOLKLEKAPKKGIT-AMVKEVILQSLVUDGNY 60 Db 158 KTKNSDDEBLKQKQLELEKVK---IDQTADEKKGITREALKBETVKNSG-- 210 Qy 61 DCERIGTSNYYWAFPS-KALHARHKLVEU---BSQLEBGSOKRASLQKST----- 107 Db 211 ---LSTTSELALAKTQVLSLEKEKESLQFLSGNKSKELEDLYIQHSDISEKUKALTDL 265 Qy 108 EKAKIGRCETEERTRLAKELLSRDRQLEKAEVYKODDQVWVERIRQANKVKEAA 166 Db 267 KEKIQFDFDSKQKLTELENDLSTKCELETEKTOTSKENLERKDEIVKLNKHELEK 326 Db 327 NT---DN---SGAKKELLERVKSL 344

Query Match Best Local Similarity 10.6%; Score 111; DB 4; Length 1937; Matches 54; Conservative 42; Mismatches 88; Indels 50; Gaps 9; Qy 4 KKGGLSABERKRTRMMEIFSETKDVFOLKLEKAPK-----EKGITAMS-----V 47 Db 1361 QRALSKANSEVQWQRTYEDAIORTEELEBAAKLAQRLQBAEHEVRAUNAKASLRT 1420 Qy 48 KEVLOSLVUDGNDGMDRCERTIGTS-----NYYWAFPSKALHARHKLVEQSLRSQ 98 Db 1421 KORLQNEVEDMLMVERSNAACALDKQRF-----DKVLSERWQKQFETQAELESQ 1475 Qy 99 KHASLQSKLSIRAK-----IGCET-----ERTRLAKELSSRDO-----REQLKAE 140 Db 1476 ESRSLSTELFKVQVYVEESDQLETTURRNQLOQEISLTLTQIAEGKQTHELMKIKQ 1535 Qy 141 VEKKKODDQVWVERIRQANKVKEAA-ANRWTNIFAIKSWATRKEFQEENKIDR 1535 Db 1536 VEQK-CEIQAEELEBASLHERSKLIRQLDINQNOVSEVDRKIKBEBEID 1588

RESULT 11
 US-09-914-259-46
 ; Sequence 46, Application US/09914259
 ; Patent No. 6495336
 ; GENERAL INFORMATION:
 ; APPLICANT: Makowski, Lee
 ; APPLICANT: Hyman, Paul
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
 ; FILE REFERENCE: 8471-010-939
 ; CURRENT APPLICATION NUMBER: US/09/914,259
 ; NUMBER OF SEQ ID NOS: 180
 ; SEQ ID NO: 46
 ; LENGTH: 284
 ; TYPE: PRT
 ; ORGANISM: Brachydanio rerio
 ; US-09-914-259-46

Query Match Best Local Similarity 10.5%; Score 110; DB 4; Length 284; Matches 51; Conservative 44; Mismatches 72; Indels 52; Qy 108 EKAKIGRCETEERTRLAKELLSRDRQLEKAEVYKODDQVWVERIRQANKVKEAA 176 Db 177 ER-----TERRAELAESKOSLEELKEVNTUNLKSLEAQAKYISQKEDQYHEEIKLT 159 Qy 160 KVAKELANRWTNIFAIKSWATRKEFQEENKIDR 229 Db 230 DKLKEAETRAE--FAERSVA-----KLEKTIIDDED 258

RESULT 10
 US-09-538-092-918

Page 6

TITLE: REILICH WEINSTOCK et al
TITLE OF INVENTION: NUCLEAR ACTIVATION

1 CURRENT REFERENCE: 107196_132
1 CURRENT APPLICATION NUMBER: US/09/248,796A
1 CURRENT FILING DATE: 1999-02-12
1 PRIORITY NUMBER: US 60/074,725
1 PRIORITY FILING DATE: 1998-02-13
1 PRIORITY APPLICATION NUMBER: US 60/096,409
1 NUMBER OF SEQ ID NOS: 28208
1 S30 ID NO: 20276
1 LENGTH: 817
1 TYPE: PRT
1 ORGANISM: *Candida albicans*
US-09-248-796A-20276

```

          10.3%; Score 108; DB 4; Length 817;
          best local Similarity 24.4%; Pred. No. 0.059;
          Matches 49; Conservative 43; Mismatches 85; Indels 24; Gaps 6
OY 9 AEBKTRMMEIFSETCDVOLKDKIETKAPKEKGITANMSVK--EVQSLVDP---DCWVCE 63
Db 426 AEEKVQ--LLODEKKENMNDRIDTEAKIAKARBOSLELQAEDKELPRTJUDLELKESAKLE 484
QY 64 RIGTNYYWAFPSKLUH---ARKHKFUVLESQSLSGQSOKHASLQSIKAKGRCRTEEE 119
Db 485 EUTNRADELANEVYQASEDLNKEYEVEKLAELSKLQBKAONDIKEYCTTDEA----TAK 538
QY 120 RTRLIKKESSLRDOREBOLKAEVK YKDCPQVWVTEIROQNKVAEANRWTNDIIFAKS 178
Db 539 HESTDKVEAELQBLHKEDKADAEKEHEDLQGKSLEROKQKHLDEDDAKTKKEDILAAIDE 598
QY 179 WAKRKGKEBENKLDRFGIE 199
Db 599 KVK-----DEHKINSE--LKE 612

```

GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: March 23, 2005, 19:17:58 ; Search time 138 seconds
(without alignments)
491.853 Million cell updates/sec

Title: US-10-087-190-3
Perfect score: 1047
Sequence: 1 MSKKKGISAEKKRTRMMEIF.....FEENKIDRTFGIPBDFDHYID 205
Scoring table: Blosum62
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 33110923 residues
Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgm2_6/prodata/1/pubpa/US07_PUBCOMB.pep:*

2: /cgm2_6/prodata/1/pubpa/FCT1_NEW_PUB.pep:*

3: /cgm2_6/prodata/1/pubpa/US06_PUBCOMB.pep:*

4: /cgm2_6/prodata/1/pubpa/US07_NEW_PUB.pep:*

5: /cgm2_6/prodata/1/pubpa/US07_PUBCOMB.pep:*

6: /cgm2_6/prodata/1/pubpa/PCTUS_PUBCOMB.pep:*

7: /cgm2_6/prodata/1/pubpa/US08_PUBCOMB.pep:*

8: /cgm2_6/prodata/1/pubpa/US09_PUBCOMB.pep:*

9: /cgm2_6/prodata/1/pubpa/US09A_PUBCOMB.pep:*

10: /cgm2_6/prodata/1/pubpa/US09B_PUBCOMB.pep:*

11: /cgm2_6/prodata/1/pubpa/US09C_PUBCOMB.pep:*

12: /cgm2_6/prodata/1/pubpa/US09_NEW_PUB.pep:*

13: /cgm2_6/prodata/1/pubpa/US10_PUBCOMB.pep:*

14: /cgm2_6/prodata/1/pubpa/US10B_PUBCOMB.pep:*

15: /cgm2_6/prodata/1/pubpa/US10C_PUBCOMB.pep:*

16: /cgm2_6/prodata/1/pubpa/US10D_PUBCOMB.pep:*

17: /cgm2_6/prodata/1/pubpa/US10_NEW_PUB.pep:*

18: /cgm2_6/prodata/1/pubpa/US11_NEW_PUB.pep:*

19: /cgm2_6/prodata/1/pubpa/US60_NEW_PUB.pep:*

20: /cgm2_6/prodata/1/pubpa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1047	100.0	205	Sequence 2, Appli
2	1047	100.0	205	Sequence 3, Appli
3	1047	100.0	205	Sequence 14, Appli
4	1047	100.0	205	Sequence 20, Appli
5	1047	100.0	205	Sequence 21, Appli
6	1047	100.0	205	Sequence 22, Appli
7	1047	100.0	205	Sequence 61, Appli
8	1047	100.0	205	Sequence 1821, Appli
9	1036.5	99.0	206	Sequence 44, Appli
10	1036.5	99.0	206	Sequence 45, Appli
11	1011	98.6	198	Sequence 24, Appli
12	975	93.1	190	Sequence 13, Appli
13	975	93.1	190	Sequence 19, Appli

RESULT 1
US-09-799-250-2
; Sequence 2, Application US/09799250
; Publication No. US20030032087A1
; GENERAL INFORMATION:
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Rene S. Hubert
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Paris
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129_3.US01
; CURRENT APPLICATION NUMBER: US/09/799,250
; NUMBER OF SEQ ID NOS: 719
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-250-2
Query Match 100.0%; Score 1047; DB 10; Length 205;
Best Local Similarity 100.0%; Pred. No. 3e-78; DB ID: 205;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKKKGISAEKKRTRMMEIFSETKDFOLKOLEKAPKKGITAMSVKEVQSLVDDGMV 60
1 MSKKKGISAEKKRTRMMEIFSETKDFOLKOLEKAPKKGITAMSVKEVQSLVDDGMV 60
Qy 61 DCERIGTSNYYWAFSKALHARKHKLVEVLSOLSTGSKQHASLOKSIKAKIGRCETEE 120
1 DCERIGTSNYYWAFSKALHARKHKLVEVLSOLSTGSKQHASLOKSIKAKIGRCETEE 120

QY

Db

QY

RESULT 2
 US-10-087-190-3
 ; Sequence 3, Application US/10087190
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Afar, Daniel E. H.
 ; APPLICANT: Ge. Wangmao
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; CURRENT APPLICATION NUMBER: US 10/087,190
 ; CURRENT FILING DATE: 2003-01-28
 ; PRIORITY NUMBER: US 09/779,250
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 205
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-087-190-3

Query Match
 Best Local Similarity 100.0%; Score 1047; DB 15; Length 205;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKKGLSAAEKRTRMMEIFSETKDVFOLDKOLEKAPKEKGITAMSVKEVTLQSLVDDGMV 60
 Db 1 MSKKGLSAAEKRTRMMEIFSETKDVFOLDKOLEKAPKEKGITAMSVKEVTLQSLVDDGMV 60
 QY 61 DCERIGTSNYYWAFPSKALHARKHKLVELESOLSEGSKHSLQSKSTEAKIGRCSTEER 120
 Db 61 DCERIGTSNYYWAFPSKALHARKHKLVELESOLSEGSKHSLQSKSTEAKIGRCSTEER 120
 QY 121 TRIAKELSSLRDQBLKAEVEKYKCDPQWVBIROANKVAKAEANRWTNIAIKSWA 180
 Db 121 TRIAKELSSLRDQBLKAEVEKYKCDPQWVBIROANKVAKAEANRWTNIAIKSWA 180
 QY 181 KRKGFEENKIDRTFGIPEDFDYID 205
 Db 181 KRKGFEENKIDRTFGIPEDFDYID 205
 ; US-10-087-190-3

RESULT 3
 US-10-087-190-14
 ; Sequence 14, Application US/10087190
 ; Publication No. US20030223997A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys, Inc.
 ; APPLICANT: Hubert, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Afar, Daniel E. H.
 ; APPLICANT: Ge. Wangmao
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; FILE REFERENCE: 5158-2003.20
 ; CURRENT APPLICATION NUMBER: US 10/087,190
 ; CURRENT FILING DATE: 2003-01-28
 ; PRIORITY NUMBER: US 09/779,250
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 205
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-087-190-20

Query Match
 Best Local Similarity 100.0%; Score 1047; DB 15; Length 205;
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKKGLSAAEKRTRMMEIFSETKDVFOLDKOLEKAPKEKGITAMSVKEVTLQSLVDDGMV 60
 Db 1 MSKKGLSAAEKRTRMMEIFSETKDVFOLDKOLEKAPKEKGITAMSVKEVTLQSLVDDGMV 60
 QY 61 DCERIGTSNYYWAFPSKALHARKHKLVELESOLSEGSKHSLQSKSTEAKIGRCSTEER 120

RESULT 5
US-10-087-190-21
; Sequence 21, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Challita-Eld, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Marty
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmiao
; APPLICANT: Jakobovitz, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 121PI1 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20034-20
; CURRENT APPLICATION NUMBER: US/10/087-190
; CURRENT FILING DATE: 2003-01-28
; PRIORITY APPLICATION NUMBER: US 09/779,250
; PRIORITY FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 22
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-21
Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3e-78; 0; Mismatches 205; Conservative 0; Indels 0; Gaps 0;
Matches 205; Conservat 0; Mismatches 205;
Qy 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Db 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Qy 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Db 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Qy 121 TRIAKELSSLRDQREOKAEVYKODPQVEERQANKYAKEAANRWTNIAFKSWA 180
Db 121 TRIAKELSSLRDQREOKAEVYKODPQVEERQANKYAKEAANRWTNIAFKSWA 180
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 21
LENGTH: 205
TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-21
Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3e-78; 0; Mismatches 205; Conservative 0; Indels 0; Gaps 0;
Matches 205; Conservat 0; Mismatches 205;
Qy 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Db 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Qy 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Db 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Qy 121 TRIAKELSSLRDQREOKAEVYKODPQVEERQANKYAKEAANRWTNIAFKSWA 180
Db 121 TRIAKELSSLRDQREOKAEVYKODPQVEERQANKYAKEAANRWTNIAFKSWA 180
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 21
LENGTH: 205
TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-22
Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3e-78; 0; Mismatches 205; Conservative 0; Indels 0; Gaps 0;
Matches 205; Conservat 0; Mismatches 205;
Qy 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Db 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Qy 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Db 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Qy 121 TRIAKELSSLRDQREOKAEVYKODPQVEERQANKYAKEAANRWTNIAFKSWA 180
Db 121 TRIAKELSSLRDQREOKAEVYKODPQVEERQANKYAKEAANRWTNIAFKSWA 180
PRIOR APPLICATION NUMBER: US 09/779,250
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 69
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 21
LENGTH: 205
TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-22
Query Match 100.0%; Score 1047; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3e-78; 0; Mismatches 205; Conservative 0; Indels 0; Gaps 0;
Matches 205; Conservat 0; Mismatches 205;
Qy 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Db 1 MSKKGGLSAEERTRMMEIFSETKDVQLKOLEKAPKEKGITAMSKEVQLQSLVUDGMV 60
Qy 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120
Db 61 DCERIGTSNYWAFPSKALHARKHKLVELESQSEGSKQHASLQKSIEKAKIGRCBTER 120

RESULT 8
US-10-408-765A-1821
; Sequence 1821, Application US/10408765A.
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088455
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; NUMBER OF SEQ ID NO: 205
; SEQ ID NO: 1821
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1821
Query Match
Best Local Similarity 100.0%; Score 1047; DB 16; Length 205;
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSKKKGSAEKKRTRMMIFSETKDFQQLKOLEKTAPKKGITAMSVKEVLQSLVLDGMV 60
Db 1 MSKKKGSAEKKRTRMMIFSETKDFQQLKOLEKTAPKKGITAMSVKEVLQSLVLDGMV 60
Qy 61 DCERIGTSNYWAFPSKALHARKHKLVEVLSQELSEGSKHSLQSKIEKAKIGRCTEE 119
Db 61 DCERIGTSNYWAFPSKALHARKHKLVEVLSQELSEGSKHSLQSKIEKAKIGRCTEE 119
Qy 120 RTRLAKESSLRDOREOLKAESKEVYKCDPQVVEIRQANKVAKEAANRWTDNIFAKSW 179
Db 121 RTRLAKESSLRDOREOLKAESKEVYKCDPQVVEIRQANKVAKEAANRWTDNIFAKSW 179
Qy 180 AKRKGFFENKIDRTGIPEDFDYID 205
Db 181 AKRKGFFENKIDRTGIPEDFDYID 205
; Sequence 44, Application US/10087190
; Publication No. US20030223997A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Chalilta-Eid, Pia M.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E. H.
; APPLICANT: Ge, Wangmao
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE REFERENCE: 51158-20034-20
; CURRENT APPLICATION NUMBER: US/10/087, 190
; CURRENT FILING DATE: 2003-01-28
; PRIORITY APPLICATION NUMBER: US 09/779, 250
; PRIORITY FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO: 45
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-087-190-44
Query Match
Best Local Similarity 99.0%; Score 1036.5; DB 15; Length 206;
Matches 205; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 MSKKKGSAEKKRTRMMIFSETKDFQQLKOLEKTAPKKGITAMSVKEVLQSLVLDGMV 60
Db 1 MSKKKGSAEKKRTRMMIFSETKDFQQLKOLEKTAPKKGITAMSVKEVLQSLVLDGMV 60
Qy 61 DCERIGTSNYWAFPSKALHARKHKLVEVLSQELSEGSKHSLQSKIEKAKIGRCTEE 119
Db 61 DCERIGTSNYWAFPSKALHARKHKLVEVLSQELSEGSKHSLQSKIEKAKIGRCTEE 119

Db 61 SKALHARKHKLVELSQISSEGSKHSLQSKIEAKIGRCETEERTRIAKELSSLRDORE 120
 Qy 136 QLKAEVEKYKDCDPQVVEIROANKVAKAEANRWTNDNAIKSWAKRKGFFENKIDRTF 195
 Db 121 QLKAEVEKYKDCDPQVVEIROANKVAKAEANRWTNDNAIKSWAKRKGFFENKIDRTF 180
 Qy 196 GIPEDFDYID 205
 Db 181 GIPEDFDYID 190

RESULT 14
 US-10-087-190-67
 ; Sequence 67, Application US/10087190
 ; General Information: Publication No. US20030223997A1
 ; Applicant: Agensys, Inc.
 ; Applicant: Charlotte-Eid, Pia M.
 ; Applicant: Hubert, Rene S.
 ; Applicant: Raitano, Arthur B.
 ; Applicant: Faris, Mary
 ; Applicant: Afar, Daniel E. H.
 ; Applicant: Ge, Wangmao
 ; Applicant: Jakobowitz, Aya
 ; Title of Invention: NUCLEIC ACID AND CORRESPONDING PROTEIN
 ; File Reference: 51158-2004.20
 ; Current Application Number: US/10/087-190
 ; Prior Filing Date: 2003-01-28
 ; Prior Application Number: US 09/779, 250
 ; Number of SEQ ID NOS: 69
 ; Software: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 68
 ; LENGTH: 190
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 ; US-10-087-190-68

Query Match
 Best Local Similarity 93.1%; Score 975; DB 15; Length 190;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 16 MMEIFSETKDVFOLKDKLEKIAPEKGITAMSKVKEVLQSLVUDGMVDCERIGTSNYWAPP 75
 Db 1 MMEIFSETKDVFOLKDKLEKIAPEKGITAMSKVKEVLQSLVUDGMVDCERIGTSNYWAPP 75
 Qy 76 SKALHARKHKLVELSQISSEGSKHSLQSKIEAKIGRCETEERTRIAKELSSLRDORE 135
 Db 61 SKALHARKHKLVELSQISSEGSKHSLQSKIEAKIGRCETEERTRIAKELSSLRDORE 120
 Qy 136 QLKAEVEKYKDCDPQVVEIROANKVAKAEANRWTNDNAIKSWAKRKGFFENKIDRTF 195
 Db 121 QLKAEVEKYKDCDPQVVEIROANKVAKAEANRWTNDNAIKSWAKRKGFFENKIDRTF 180
 Qy 196 GIPEDFDYID 205
 Db 181 GIPEDFDYID 190

RESULT 15
 US-10-087-190-68
 ; Sequence 68, Application US/10087190
 ; General Information: Publication No. US20030223997A1
 ; General Information:
 ; Applicant: Agensys, Inc.
 ; Applicant: Charlotte-Eid, Pia M.
 ; Applicant: Hubert, Rene S.
 ; Applicant: Raitano, Arthur B.
 ; Applicant: Faris, Mary
 ; Applicant: Afar, Daniel E. H.
 ; Applicant: Ge, Wangmao
 ; Applicant: Jakobowitz, Aya

Search completed: March 23, 2005, 19:30:49
 Job time : 139 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.
 OM protein - protein search, using sw model

Run on: March 23, 2005, 19:08:37 ; Search time 41 Seconds
 (without alignments)
 481.084 Million cell updates/sec

Title: Perfect score: US-10-087-190-3
 Sequence: 1 MSKKKLSAEEKRTRMMBIF.....FENKIDRTFGIPERDFID 205
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR 79.4
 1: piri:
 2: pir2:
 3: pir3:
 4: pir4:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	349	33.3	T37610	hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
2	197	18.8	128	hypothetical prote
3	168.5	16.1	174	hypothetical prote
4	135.5	12.9	1281	dynactin 1 - mouse
5	130.5	12.5	1053	dynactin 1 - chicken
6	121.5	11.6	1356	kinectin 1 - human
7	117	11.1	768	hypothetical prote
8	116.5	11.2	2442	centrosome aspacia
9	115.5	11.0	880	conserved hypothetical
10	114	10.9	284	tropomyosin alpha, hypotethetical prote
11	114	9.4	764	tropomyosin alpha, hypotethetical prote
12	113	10.8	199	A32183
13	112	10.7	284	JC6199
14	112	10.7	285	A24199
15	112	10.7	559	S49143
16	112	10.7	1938	A59293
17	111	10.6	559	A45620
18	111	10.6	1937	T38055
19	110.5	10.6	308	T08796
20	110.5	10.6	629	T44607
21	110.5	10.6	879	C71083
22	110	10.5	284	I51731
23	109.5	10.5	168	G86578
24	109.5	10.5	168	B72046
25	109.5	10.5	284	JC6198
26	109.5	10.5	400	E70318
27	109.5	10.5	1298	T24480
28	109.5	10.5	1390	S51564
29	109	10.4	284	JC2551

RESULT 1

Query Match 33.3%; Score 349; DB 2: Length 210;
 Best Local Similarity 41.6%; Pred. No. 3.9e-16; Mismatches 76; Indels 6; Gaps 5;
 Matches 84; Conservative 36;

Qy	5	KGLSABEKKRTRMMIFMAIFSETKDVFOLKDIKEIKAPEKEKGITAMSTKENVKLSVLSNDGMDCER 64
Db	4	KGLSLAKERKRLEAFHDSKDFOLKEVEKGSK-KQIVLQTYKDVLSLVDNPKTEK 62
Qy	65	IGTSNTYWWFPPSKALHARKEKLYTESOSEGSKQKHSIQLKST--EKAKIGRCE-EKT 121
Db	63	IGTSNTYWWFPPSKALHARKEKLYTESOSEGSKQKHSIQLKST--EKAKIGRCE-EKT 122
Qy	122	RKAKELSSLRD-QREQLKAEVKEVKYDQPCPVQVISBIRQANKVAEARNWTDNPAIKWA 160
Db	123	QYTTLELLHAKESKEIKLKQIQLSINHCPETFLKNEYTKKYMHEANLWTDQIHTLIAFC 182
Qy	181	KRKFGFEEENKIDRTFGIPERDFD 202
Db	183	-RDNGADTNQIREYCSIPBDLD 203

RESULT 2

Query Match 33.3%; Score 349; DB 2: Length 210;
 Best Local Similarity 41.6%; Pred. No. 3.9e-16; Mismatches 76; Indels 6; Gaps 5;
 Matches 84; Conservative 36;

Qy	5	KGLSABEKKRTRMMIFMAIFSETKDVFOLKDIKEIKAPEKEKGITAMSTKENVKLSVLSNDGMDCER 64
Db	4	KGLSLAKERKRLEAFHDSKDFOLKEVEKGSK-KQIVLQTYKDVLSLVDNPKTEK 62
Qy	65	IGTSNTYWWFPPSKALHARKEKLYTESOSEGSKQKHSIQLKST--EKAKIGRCE-EKT 121
Db	63	IGTSNTYWWFPPSKALHARKEKLYTESOSEGSKQKHSIQLKST--EKAKIGRCE-EKT 122
Qy	122	RKAKELSSLRD-QREQLKAEVKEVKYDQPCPVQVISBIRQANKVAEARNWTDNPAIKWA 160
Db	123	QYTTLELLHAKESKEIKLKQIQLSINHCPETFLKNEYTKKYMHEANLWTDQIHTLIAFC 182
Qy	181	KRKFGFEEENKIDRTFGIPERDFD 202
Db	183	-RDNGADTNQIREYCSIPBDLD 203

hypothetical protein F19B15-200 - *Arabidopsis thaliana* (mouse-ear cress)
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 11-Jun-1999 #sequence revision 11-Jun-1999 #text_change 09-Jul-2004
 C;accession: T08972
 R;Van, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, A.;reference number: ZU6519
 A;Accession: T08972
 A;Molecule type: DNA

tropomyosin 1, alpha
 tropomyosin 1, heavy chain
 tropomyosin 2, alpha
 tropomyosin 2, fib
 tropomyosin alpha
 tropomyosin 2, fib
 myosin II heavy chain
 myosin II antigen PF
 RESA-H3

RESULT 6

S32763 ANKVAKEAANRWTDDIFAKSWAKKKGFFGKPFEDDF 202
 kinase, 1 - human
 C;Species: Homo sapiens (man)
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C;Accession: S32763; I37977
 R;Krupp, G.; Fuetterer, A.; Lemke, H.; Kroenke, M.
 submitted to the EMBL Data Library, April 1993
 A;Description: Cloning and characterization of TAF, a novel transactivating protein.
 A;Reference number: S32763
 A;Accession: S32763
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1356 <KRU>
 A;Cross-references: UNIPROT:Q86UP2; EMBL:Z22551
 R;Futterer, A.; Krupp, G.; Kramer, B.; Lemke, H.; Kroenke, M.
 Mol. Biol. Cell 6, 161-170, 1995
 A;Title: Molecular cloning and characterization of human kinase.
 A;Reference number: I37947; NID:95306853; PMID:787243
 A;Accession: I37947
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1356 <RES>
 A;Cross-references: EMBL:Z22551; NID:9296163; PID:CAA80271.1; PID:9296164
 C;Genetics:
 A;Gene: GDB:KTN1
 A;Cross-references: GDB:6165852; OMIM:600653

Query Match 11.6%; Score 121.5; DB 2; Length 1356;
 Best Local Similarity 22.7%; Pred. No. 2.3;
 Matches 50; Conservative 31; Mismatches 88; Indels 51; Gaps 5;
 Qy 3 KKKKGSAEERTRRMEIFBFTKDVOLKOLDEKIAKREKGITA--MSVKEVLSQSLDDGMV 60
 Db 1025 RKQKNDLREKWNEMALASTERKQI-QDKVNTKSKERQQVAEVEALEKEVKKLPPKSV 1084
 Qy 61 DCERIGTSVYWAFFSKALH----ARKHKEVLETSQSLSGSQKHSQLSKIEKAKIGRC 115
 Db 1085 P-SNLISYGEWLHGPEKKAKEBCMAGSTGSEEVKVLHKLKEADEMHTLQLQBCEKYKSVL 1143
 Qy 116 ETE--ERTRIA-KEISSLRD 132
 Db 1144 ETEGLIQKLUORSVQEQENKVKVQDESHKTIKQMQSSFTSSEQEIERLRSBNKDENRR 1203
 Qy 133 QRSQKIAEYKQCDPQVVEIQQANKVAKEAANRWTDN 172
 Db 1204 EREHLEMELEKAEMBRSTVYTVTEVRBLKDLTBLOKRLKUDS 1243

RESULT 7

T02572 hypothetical protein At2g39300 [imported] - Arabidopsis thaliana
 N;Alternative names: hypothetical protein T16B24.6
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Accession: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
 C;Accession: T02572; F84815
 R;Rounseley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, A.; Submitted to the BMBL Data Library, August 1998
 A;Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.
 A;Reference number: ZL4679
 A;Accession: T02572
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1768 <ROU>
 A;Cross-references: UNIPROT:Q80951; EMBL:AC04697; NID:93402671; PID:93402677
 A;Experimental source: cultivar Columbia
 R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;

eubs, D.; Nieman, W.C.; white, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; *Nature* 402, 761-768, 1999
 Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A; Reference number: A84420; MNUID:20083487; PMID:10617197
 A; Accession: P84815
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-768 <STO>
 A; Cross-References: GB:AE002993; NID:93402677; PIDN: AAC28980.1; GSPDB:GN00139
 C; Genetics:
 A; Gene: T16B4.6; At2g39300
 A; Map position: 2
 A; Introns: 80/2; 665/3

Query Match 11.2%; Score 117; DB 2; Length 768;
 Best Local Similarity 22.1%; Pred. No. 2.5;
 Matches 50; Conservative 46; Mismatches 90; Indels 40; Gaps 9;

QY	1	MSKKGGLSAB-----EKRTRMMEIFSETK-----DVFOI-KDLEKIAPKEGKIGT	43
Db	247	MCKKEDDVSSLELKRYKEAEGRKVLSEEMEEKKFSDPDISYLGDIQMEBERVGL-	305
QY	44	AMSVEVRLQSLVDPGQMVNDCEIRGHTSYNNWAFPSKHLHARKHKLVEY-LESOLOSEGSKRHAS	102
Db	306	AFEVVLISLRLSQMDERASTBDRDIRVKNWDMLLKKRLEKEKTYLQVQLETELDRRSSEWTS	365
QY	103	LQKSEIKAIGRCETEERLAKELSSLRQDREQKJAEVKQDQPVVEIQRQANKVA	162
Db	366	---KESFKY-----EERKRLRERRELAKHNVSQRE-ESTPHEKTERIDMIRHLDETV	416
QY	163	KE----AANRWTDNIAFKSAWAKRKGFFEENKIDRTFGIPEDDFDXI	204
Db	417	TELSATAEEMREENFLMQNLSKQESYT-----GSTDLDYV	454

A;Reference number: S623997
 A;Accession: S63011
 A;Molecule type: DNA
 A;Residues: 1-199 <POW>
 A;Cross-references: EMBL:271355; NID:91301970; PIDN:CAA95953.1; PID:91301971; MIPS:YNL07
 A;Experimental source: strain S288C
 R;Soler-Mira, A.; Saiz, J.E.; Ballesta, J.P.G.; Remacha, M.
 submitted to the Protein Sequence Database, April 1996
 A;Reference number: S63018
 A;Accession: S63018
 A;Residues: 1-199 <SOL>
 A;Cross-references: EMBL:271355; NID:91301970; PIDN:CAA95953.1; PID:91301971; MIPS:YNL07
 A;Experimental source: strain S288C
 R;Pohlmann, R.; Philippson, P.
 Yeast 12, 391-402, 1996
 A;Title: Sequencing of a cosmid clone of *Saccharomyces cerevisiae* chromosome XIV reveals 12
 A;Reference number: S63925; MUID:9626764; PMID:8701611
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-199 <POF>
 A;Cross-references: EMBL:X86470; NID:9791101; PIDN:CAA60179.1; PID:9791105
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
 C;Genetics:
 A;Gene: SGD:TPM1
 A;Cross-references: SGD:S0005023; MIPS:YNL079C
 A;Map position: 14L
 A;Superfamily: tropomyosin TPM1
 C;Keywords: coiled coil; cytoskeleton

 Query Match 10.8%; Score 113; DB 2; Length 199;
 Best Local Similarity 24.7%; Pred. No. 1;
 Matches 46; Conservative 40; Mismatches 54; Indels 46; Gaps 8;
 QY
 7 LSAEKKRTRMELFSETKQVQFQLKDLKAPKKEKQITAMSVKEVITQSLVDD-----GMV 60
 13 LEAESWQERKVEELKEKQKLEQ-ENV-----KENQIKSLTVKN-----QOLEBIEKIEAGLS 65
 OY
 61 DCERIGTSNYYWAFFSKALHARKHKL---EVLESQSEG-----SOK 99
 66 DSKTEQDNKEVQKNSITVKNHQLBEETEKLKAEKQKLSQDSHHTQSNNNNFSKK 125
 100 HASLQKSTIEKAKIGRCETERTR-----LAKELSSLRDQEQIKAEVB---KVKDC 147
 126 NQOLEEDLERSDTPKETTEKLRSSDLQADQLERVVALEBQREWERKNEBELTVKVEDA 185
 QY 148 DPOVVE 153
 Db 186 KKEELDE 191

 RESULT 13
 JC6199
 alpha-tropomyosin S-1 - axolotl
 ;Species: *Ambystoma mexicanum* (axolotl)
 C;Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
 C;Accession: JC6199
 R;Liuque, E.A.; Spinner, B.J.; Dube, S.; Dube, D.K.; Lemanski, L.F.
 Gene 185, 175-180, 1997
 A;Title: Differential expression of a novel isoform of alpha-tropomyosin in cardiac and
 A;Reference number: JG198; MUID:97208870; PMID:9055812
 A;Contents: skeletal muscle
 A;Accession: JC6199
 A;Molecule type: mRNA
 A;Residues: 1-284 <LIUQ>
 A;Cross-references: UNIPROT:P87349; GB:U33450; NID:91871357; PIDN: AAC60092.1; PID:918713
 C;Comment: This protein is a actin-binding protein.
 C;Genetics:
 C;Superfamily: tropomyosin
 C;Keywords: actin binding

Query	Match	Score	DB	Length
Best	Local Similarity	22.7%	DB 2;	Length 284;
Matches	48;	Conservative	41;	Pred. No. 1.8;
			54;	Mismatches 74;
			48;	Indels 9;
			48;	Gaps 9;
Oy	1 MSKKKGLSAEE--KTRMMBIFSETKDVFO-----LKDLERKAPK-EKGTTAMSV	47		
Db	73 LARKKATDAESDVASINRRIOLVEELDRAQERLATALQKDEEAKADESERGMVKIEN	132		
Oy	48 KEVLOSLVDDGWDGDCERIGTSNYWAPPSKALHARKHKLVELESQLSSESQK---HASL	103		
Db	133 R-----ALKDEKELMQLQEI-----QLOPERAKHIAEBADRKEEVARKLVILIEGDL	176		
Oy	104 OKSIEHKAKI--GRCTEETRFLAKELISSLDORESOLKAIEVKEVYKDDQPVVEETRANKV	161		
Db	177 ERAEERBELSEGKC-----AELEEBLKTVINNLKSLLEAQAEKTSOKEDKYEEIKVLTDK	231		
Oy	162 AKEAARWTDNIFAIKSWAKRGKFGEENKID	192		
Db	232 LKEAER-----AFAEARTVAKLEKSTD	254		

CIDate: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 15-Mar-2004

CAccession: S49143

R;Frosch, P.M.; Hartmann, M.; Sygulla, L.; Margutti, P.; Frosch, M.

Submitted to the EMBL Data Library, January 1994

A;Description: Identification of a cDNA clone from the larval stage of *Echinococcus granulosus*

A;Accession: S49143

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-559 <FRO>

C;Superfamily: EMBL:229489; NID:9509759; PID:CAA82625.1; PID:9509760

F;12-293/Domain: protein 4.1 membrane-binding domain homology <B41>

Query Match Score 112; DB 2; Length 559;

Best Local Similarity 10.7%; Pred. No. 3.7;

Matches 49; Conservative 34; Mismatches 61; Indels 32; Gaps 7;

Qy 1 MSKKKGISAEKRTRMEEIFSETKVQFQDLEKIAPKRGITAMSVKVLSVLDQMV 60

Db 294 MRRKSDSIEVQMKI----QAKSERELKEAEQRQLKEERLORMENBQLRRAQ-MV 347

Qy 61 DCERIGTSNYWAWFFSKALMARGKHLVEFQSLQSE-----GSQKHA--SLOQSIK-AKI 112

Db 348 EKE-----SDIADMKNGASAYESKIALEMILQOQRHARBSLSOKSOKLAE 394

Qy 113 GRCETEERTRLAKELSSLRQREOKAETEKYKCDPQWEEFRQANKVAKGAANR 168

Db 395 NRKLUKEETAAASAERNRLLMQRDVEOREVEAQK-----VAMAKKEAERKAQARL 445

Search completed: March 23, 2005, 19:18:39
Job time : 43 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.									
OM protein - protein search, using SW model									
Run on: March 23, 2005, 19:08:01 ; Search time 177 Seconds (Without alignments)									
593.086 Million cell updates/sec									
Title: US-10-087-190-3									
Perfect score: 1047									
Sequence: 1 MSKKKGISAEKKTRMMEIF.....FRENKIDRTFGKIPEDFDYID 205									
Scoring table: BLOSUM62									
Searched: Gapop 10.0 , Gapext 0.5									
Total number of hits satisfying chosen parameters: 1612378									
Minimum DB seq length: 0									
Maximum DB seq length: 200000000									
Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries									
Database : UniProt 03:*									
1: uniprot_sprot:*									
2: uniprot_trembl:*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result and is derived by analysis of the total score distribution.									
SUMMARIES									
Result	Score	Query	%	Match	Length	DB	ID	Description	RP
1	1047	100.0	205	2	Q9BWT6	Q9BWT6	Q9BWT6	RESULT 1	SEQUENCE FROM N.A.
2	954	91.1	205	2	Q9K396	Q9BWT6	Q9BWT6	PRELIMINARY;	RP
3	948	90.5	205	2	Q9D0A1	Q9BWT6	Q9BWT6	PRT; 205 AA.	RC
4	644	61.5	220	2	Q6DC61	Q9BWT6	Q9BWT6	DT	TIEMBLrel.
5	462	44.1	196	2	Q6E6Z8	Q9BWT6	Q9BWT6	DT	17, last sequence update)
6	440	42.0	230	2	Q8GYD2	Q9BWT6	Q9BWT6	DT	25-OCT-2004 (TIEMBLrel. 28, last annotation update)
7	427	40.8	207	2	Q6H432	Q9BWT6	Q9BWT6	DE	GAJ.
8	349	33.3	210	1	YAS5_SCHPO	Q9BWT6	Q9BWT6	GN	Name=GAJ;
9	340.5	32.5	203	2	Q8SU9	Q9BWT6	Q9BWT6	OS	Homo sapiens (Human)
10	340.5	32.5	203	2	Q6WDA3	Q9BWT6	Q9BWT6	OC	Mammalia; Butheria; Chordata; Craniata; Vertebrata; Euteleostomi; OX
11	340.5	32.5	203	2	Q9QTX0	Q9BWT6	Q9BWT6	RL	NCBI_TAXID=9606;
12	332.5	31.8	179	2	Q6Q9F9	Q9BWT6	Q9BWT6	RN	SEQUENCE FROM N.A.
13	277	26.5	196	2	Q7RH53	Q9BWT6	Q9BWT6	[1]	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
14	266.5	25.5	211	2	Q6BQL5	Q9BWT6	Q9BWT6	[2]	TISSUE=Lymph;
15	221	21.1	225	2	Q5CH5	Q9BWT6	Q9BWT6	RR	Strainberg R.;
16	206.5	19.7	225	2	Q6CSX5	Q9BWT6	Q9BWT6	RR	Subramberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
17	197	18.8	128	2	Q6S2E5	Q9BWT6	Q9BWT6	RR	Hopkins R.D., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
18	175	16.7	222	2	Q6FL56	Q9BWT6	Q9BWT6	RR	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Cabaniss T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Brange C., Rabinowitz J., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Rhee S.S., McLellan P.J., McKernan K.J., Malek J.A., Gunnarsson P.H., Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Keeteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M., Butterfield Y.S., Rodriguez M.I., Skalska U., Smalius D.E., Schnarch A., Schein J.E., Jones S.J., Matra M.A., Krzywinski M.I., Smailius D.E., Schnarch A., Schein J.E.,
19	168.5	16.1	174	1	YGT3_YEAST	Q9BWT6	Q9BWT6	RR	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences,"
20	153.5	14.7	101	2	Q6E679	Q9BWT6	Q9BWT6	RR	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
21	136.5	13.0	1281	1	DYNA_MOUSE	Q9BWT6	Q9BWT6	RR	[3]
22	135.5	12.9	1281	1	Q6AWB1	Q9BWT6	Q9BWT6	RR	SEQUENCE FROM N.A.
23	134.5	12.8	890	2	Q6AWB3	Q9BWT6	Q9BWT6	RR	TISSUE=Lymph;
24	134.5	12.8	890	2	Q6AWB3	Q9BWT6	Q9BWT6	RR	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
25	134.5	12.8	1139	1	Q61Q37	Q9BWT6	Q9BWT6	RR	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
26	134.5	12.8	1264	2	Q6NZM3	Q9BWT6	Q9BWT6	RR	EMBL; BC02142; AAH32142.1; -.
27	134.5	12.8	1278	1	DNA_HUMAN	Q9BWT6	Q9BWT6	RR	InterPro; IPR05647; Mnd1.
28	134.5	12.8	1278	2	Q6MZ23	Q9BWT6	Q9BWT6	RR	InterPro; IPR09058; Wing_hlx_DNA_bnd.
29	131.5	12.6	1280	1	DNA_RAT	Q9BWT6	Q9BWT6	RR	PFam; PF03962; Mnd1; I.
30	130.5	12.5	1224	1	DNA_CHICK	Q9BWT6	Q9BWT6	RR	SEQUENCE FROM N.A.
31	126.5	12.1	1232	2	Q6PCJ1	Q9BWT6	Q9BWT6	RR	Query Match 100.0%; Score 1047; DB 2; Length 205;

Best Local Similarity 100%; Pred. No. 5-9e-55; Matches 205; Conservative 0; Misnatches 0; Indels 0; Gaps 0;

Qy 1 MSKKGGLAEEKTRMMEIFSETKDVFOLKDLDEKTAPEKKGITAMSVKEVLOSLVUDGMV 60
 Db 1 MSKKGGLAEEKTRMMEIFSETKDVFOLKDLDEKTAPEKKGITAMSVKEVLOSLVUDGMV 60
 Qy 61 DCERIGTSNYWAFPSKALHARKHKLVEVRSOLSESGSKHASLOKSIEKAKIGRETER 120
 Db 61 DCERIGTSNYWAFPSKALHARKHKLVEVRSOLSESGSKHASLOKSIEKAKIGRETER 120
 Qy 121 TRIAKLSSLSDRQKALEVEKVKDOPVVEIRANKVAKAANRWTNDNIAFKSWA 180
 Db 121 TRIAKLSSLSDRQKALEVEKVKDOPVVEIRANKVAKAANRWTNDNIAFKSWA 180
 Qy 181 KRKGFEENKDRTRGIPEDFDYID 205
 Db 181 KRKGFEENKDRTRGIPEDFDYID 205
 Qy 181 KRKGFEENKDRTRGIPEDFDYID 205
 Db 181 KRKGFEENKDRTRGIPEDFDYID 205

RESULT 2
 Q8K396 ID Q8K396 PRELIMINARY; PRT; 205 AA.
 AC DT 01-OCT-2002 (TREMBirel. 22, Created)
 DT 01-OCT-2002 (TREMBirel. 26, last sequence update)
 DE GAU protein.
 GN Name=2610034E181k;
 OS Mus musculus (Mouse).
 OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; OX NCBI_TaxID=10090;
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=CZCZ II; TISSUE=Mammary tumor;
 RA MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Klausner R.D., Collins F.S., Grouse L.H., Derge J.G.,
 RA Altshuller S.F., Feingold B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Blatchko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., McDonald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Bosak S.A., Loquai Lino N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia P.H.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman J., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y.,
 RA Rodriguez A.C., Grinwood J., Green E.D., Dickson M.C.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E.,
 RA Jones S.J., Marra M.A., Butterfield Y.S.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZCZ II; TISSUE=Mammary tumor;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC027741; AAH27741.1; -;
 DR MGI; MGI:92415; 2610034E181k.
 DR InterPro; IPR005647; Mnd1.1.
 DR Pfam; PF03962; Mnd1.1.
 SQ SEQUENCE FROM N.A.
 205 AA; 23849 MW; 122C3FA9B4325120 CRC64;

Query Match 91.1%; Score 954; DB 2; Length 205;
 Best Local Similarity 89.8%; Score 954; DB 2; Length 205;
 Matches 184; Conservative 10; Mismatches 11; Indels 0; Gaps 0;
 Qy 1 MSKKGGLAEEKTRMMEIFSETKDVFOLKDLDEKTAPEKKGITAMSVKEVLOSLVUDGMV 60
 Db 1 MSKKGGLAEEKTRMMEIFSETKDVFOLKDLDEKTAPEKKGITAMSVKEVLOSLVUDGMV 60
 Qy 61 DCERIGTSNYWAFPSKALHARKHKLVEVRSOLSESGSKHASLOKSIEKAKIGRETER 120
 Db 61 DCERIGTSNYWAFPSKALHARKHKLVEVRSOLSESGSKHASLOKSIEKAKIGRETER 120
 Qy 121 TRIAKLSSLSDRQKALEVEKVKDOPVVEIRANKVAKAANRWTNDNIAFKSWA 180
 Db 121 TRIAKLSSLSDRQKALEVEKVKDOPVVEIRANKVAKAANRWTNDNIAFKSWA 180
 Qy 181 KRKGFEENKDRTRGIPEDFDYID 205
 Db 181 KRKGFEENKDRTRGIPEDFDYID 205
 Qy 181 KRKGFEENKDRTRGIPEDFDYID 205
 Db 181 KRKGFEENKDRTRGIPEDFDYID 205

RESULT 3
 Q8K396 ID Q8K396 PRELIMINARY; PRT; 205 AA.
 AC DT 01-JUN-2001 (TREMBirel. 17, Created)
 DT 01-JUN-2001 (TREMBirel. 17, Last sequence update)
 DE Mus musculus 10 days embryo whole body cDNA. RIKEN full-length
 DE enriched library, clone:2610034E181b product:GAU homolog.
 GN Name=2610034E181k;
 OS Mus musculus (Mouse).
 OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; OX NCBI_TaxID=10090;
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=CZCZ II; TISSUE=Whole body;
 RA MEDLINE=9219263; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.,
 RA "High-efficiency full-length cDNA cloning.",
 RA RIKEN FANTOM Consortium;
 RA "Functional annotation of a full-length mouse cDNA collection.",
 RA Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA The FANTOM Consortium,
 RA "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RA "Analysis of the mouse genome transcriptome based on functional annotation of
 RA 60,770 full-length cDNAs.",
 RA Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayashizaki Y., Muramatsu M., Hayashizaki Y., Shibata K., Itoh M.,
 RA "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.",
 RA Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZCZ II; TISSUE=Whole body;
 RA MEDLINE=20530913; PubMed=1107861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,
 RA Kono H., Aizawa J., Nishi K., Kita K., Kita K.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashimagi A.,
 RA Fujiwake S., Inoue K., Togawa Y., Iizawa O., Ohara B., Watanuki M.,
 RA Oneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RA "RIKEN integrated sequence analysis (RISA) system-384 format

RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genome Res. 10:1757-1771(2000).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RA Adachi J., Aizawa K., Akhira S., Akimura T., Arai A., Aono H., Arai T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imotani K., Iishi Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Konno H., Kouga M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shiba Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayasizaki Y.;

RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AK011664; BABY567; DR

DR MGD: MGI:11924165; 2610034818Rik.

DR Pfam; PF03962; Mnd1; 1.

DR SEQUENCE 205 AA; 23909 MW; 09368E19B4224021 CRC64;

Query Match 90.5%; Score 948; DB 2; Length 205; Best Local Similarity 89.3%; Pred. No. 5 6e-50; Matches 183; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MSKKKGLSAEERTRMMEIFETKDVQFOLKOLEKIAPEKEGITAMSVKEVQLQSLOKSIKAKIGRCTEEER 60

Db 1 MSKKKGLSAEERTRMMEIFETKDVQFOLKOLEKIAPEKEGITAMSVKEVQLQSLOKSIKAKIGRCTEEER 60

Qy 61 DGERIGTSNYWAFPSKALHARKHKKLEVLQESQKHSLOKSIEKAKIGRCTEEER 120

Db 61 DGERIGTSNYWAFPSKALHARKHKKLEVLQESQKHSLOKSIEKAKIGRCTEEER 120

Qy 121 TRLAKELSSLADOREQKAEVYKQDQPVVEETRQANKYAKEAANRWTDNIF 180

Db 121 TRLAKELSSLADOREQKAEVYKQDQPVVEETRQANKYAKEAANRWTDNIF 180

Qy 181 AMLAKELFSFRDQROQLKAEVYKRECDPQVEEIRANKYAKEAANRWTDNIFAIKSWA 180

Db 181 AMLAKELFSFRDQROQLKAEVYKRECDPQVEEIRANKYAKEAANRWTDNIFAIKSWA 180

RESULT 4

Q6DC61 PRELIMINARY; PRT; 220 AA.

ID Q6DC61; AC 06DC61; DT 25-OCT-2004 (T-EMBLrel. 28, Created)

DT 25-OCT-2004 (T-EMBLrel. 28, last sequence update)

DE Zgc:101017 protein.

GN Name=zgc:101017;

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OX NCBI_TAXID=7955;

RA [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Singapore local strain; TISSUE=Embryo; STRAIN=12477932; DOI=10.1073/pnas.242608999;

RA Strausberg R.L., Fengold E.B., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schuler G.D., Altschul S.F., Zeeberg B.B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore S.I., Wang Y.J., Heilek F., Diatchenko L., Maruyama K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schatz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loqueline N.A., Peters G.J., Abramson R.D., Mulahay S.J., Bosak S.A., McElvan P.J., McKernan K.J., Malek J.A., Gunnarson P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fabey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shvchenko Y., Bouffard G.G., RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grinwood J., Schmutz J.J., Myers R.M., Butterfield Y.S., RA Krzywinski M.I., Skalska U., Smailus D.E., Scherich A., Schein J.E., Jones S.J., Matra M.A.;

RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; A. 99:16899-16903 (2002).

RN [12]

RP SEQUENCE FROM N.A.

RC STRAIN=Singapore local strain; TISSUE=Embryo; Director MGC Project; Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

RA EMBL; BC078223; AAH8223 1; -.

DR InterPro; IPR005647; Mnd1; 1.

DR InterPro; IPR005647; Mnd1; 1.

DR InterPro; IPR009058; Wing-hlx_DNA_brd.

DR Pfam; PF03962; Mnd1; 1.

DR SEQUENCE 220 AA; 25176 MW; 90DBAA6331F4BF7 CRC64;

Query Match 61.5%; Score 644; DB 2; Length 220; Best Local Similarity 71.8%; Pred. No. 1.4e-31; Matches 125; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MSKKKGLSAEERTRMMEIFETKDVQFOLKOLEKIAPEKEGITAMSVKEVQLQSLOKSIKAKIGRCTEEER 60

Db 1 MSKKKGLSAEERTRMMEIFETKDVQFOLKOLEKIAPEKEGITAMSVKEVQLQSLOKSIKAKIGRCTEEER 60

Qy 61 DGERIGTSNYWAFPSKALHARKHKKLEVLQESQKHSLOKSIEKAKIGRCTEEER 120

Db 61 DGERIGTSNYWAFPSKALHARKHKKLEVLQESQKHSLOKSIEKAKIGRCTEEER 120

Qy 121 TRLAKELSSLADOREQKAEVYKQDQPVVEETRQANKYAKEAANRWTDNIF 174

Db 121 EDLKLKETALKQORDQMVKEVYKQDQPAVEEIRANIAKEAVMRWTGGTF 174

RESULT 5

Q86EZB PRELIMINARY; PRT; 196 AA.

ID Q86EZB; AC 086EZB; DT 01-JUN-2003 (T-EMBLrel. 24, Created)

DT 01-OCT-2003 (T-EMBLrel. 25, last sequence update)

DE Clone ZDD1259 mRNA sequence.

OS Schistosoma japonicum (Blood fluke); Trematoda; Digenea; Strigeida; Schistosomatidae; Schistosomatida; Schistosoma.

OC Bucephala; Metzcoa; Platylminthes; Trematoda; Digenea; Strigeida; Schistosomatida; Schistosomatidae; Schistosoma.

OX NCBI_TAXID=6182;

RA [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=22879925; PubMed=12973349; DOI=10.1038/ng1236; Hu W., Yan Q., Shen D.K., Liu F., Zhu Z.D., Song H.D., Xu X.R., Wang Z.J., Rong Y.P., Zeng L.C., Wu J.J., Zhang X., Wang J.J., Xu X.N., Wang S.Y., Fu G., Zhang X.L., Wang Z.Q., Brindley P.J., McManus D.P., Xue C.L., Feng Z., Chen Z., Han Z.G.;

RA "Evolutionary and biomedical implications of a Schistosoma japonicum complementary DNA resource."; RT Nat. Genet. 35:139-147 (2003).

RA DR EMBL; AY223066; AAP06089 1; -.

DR InterPro; IPR005647; Mnd1; 1.

DR Pfam; PF03962; Mnd1; 1.

DR SEQUENCE 196 AA; 23163 MW; B30PF6F088D7123F0 CRC64;

Query Match 44.1%; Score 462; DB 2; Length 196; Best Local Similarity 47.4%; Pred. No. 1.2e-20; Matches 92; Conservative 34; Mismatches 68; Indels 0; Gaps 0;

Qy 11 EKRTTRMMEIFETKDVQFOLKOLEKIAPEKEGITAMSVKEVQLQSLOKSIKAKIGRCTEEER 70

Db 2 KSRQRMDPFKEKKDFQQLKSLERLQKKECINSVSKVDLMLVHGLUTDKGTSVY 61

Qy 71 YWAFPSKALHARKHKKLEVLQESQKHSLOKSIEKAKIGRCTEEERTRLAKELSSL 130

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Tivey A., Walsh S.V., Warren T., Whitehead S., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodhar J., Volkart G., Aert R., Robins J., Grunprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilber H., Borzyk K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Punelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelauze V., Motter S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Usary D., Barrell B.G., Nurse P.; RT "The genome sequence of *Schizosaccharomyces pombe*"; RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial/ CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: 254096; CAA90804.1; -. DR PIR: T37650; T37610; DR Genebank_Spombe; SPAC13A11.03; -. DR InterPro; IPR005647; Mnd1; DR InterPro; IPR005647; Mnd1; DR InterPro; IPR005647; HPR_Serp_S; DR InterPro; IPR005647; Wing_hlx_DNA_bnd; DR Pfam; PF03962; Mnd1; 1. DR PROSITE; PS00599; PRS_HPR_SER; UNKNOWN_1. DR KW Hypothetical protein; KW SEQUENCE 203 AA; 23743 MW; DFE833A65A1A28A42 CRC64; DR Query Match 32.5%; Score 340.5; DB 2; Length 203; DR Best Local Similarity 35.1%; Pred. No. 2.8e-13; DR Matches 71; Conservative 50; Mismatches 72; Indels 9; Gaps 3; DR QY 7 LSAEKKTRMMEIFSTKTDVFLDKLDEKIKAPKEKGITAMSVKVQSLVQSLVDDNIVTEK 66 DR 6 MKSDQKSKILLBIRGSKSFFKQELDSLGSK-KGIVVNITKEIQLQDVLGVNTRKVG 64 DR Db 67 TSNYKWAFFSKALHAKHKHLVLESQSESGSKHASLQSKTEAKIGRCETEERTRLAKE 126 DR QY 65 TSNLYNSFASIGKQCKRLCBLMECCRMSQDCIKRKEVYENEKSKHYTERNLLENK 124 DR Db 127 LSSL---RDRBQLAKEVKYKODCPQVVEIROANKVAKEAANTWDNIAFIKSWAKR 182 DR 125 LNLMKIEQDRE---ELGEFETDPIAVDKLADRKEADECRILIDNVIQDYLICS 180 DR QY 183 KPGFEENKIDRFGIPBFDYI 204 DR Db 181 KPPMEKSFNSFGIPPDLDYI 202

RESULT 9

Q8SU9 PRELIMINARY; PRT; 203 AA.

AC Q8SU9; DT 01-JUN-2002 (T-EMBLrel. 21, Created) DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update) DE Hypothetical protein; EC01_1600; Name=EC01_1600; OS Encephalitozoon cuniculi GB-M1; OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon. OX NCBI_TaxID=5741; RN [1] RP SEQUENCE FROM N. A. RA Rameesh M.A., Malik S.B., Logsdon J.M. Jr.; RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases. DR EMBL; AY285092; AAQ04512.1; -. DR InterPro; IPR005647; Mnd1; 1. DR InterPro; IPR009058; Wing_hlx_DNA_bnd; DR Pfam; PF03962; Mnd1; 1. DR SEQUENCE 203 AA; 23278 MW; CC0625DCNC4158C0 CRC64; DR Query Match 32.5%; Score 340.5; DB 2; Length 203; DR Best Local Similarity 34.8%; Pred. No. 2.8e-13; DR Matches 70; Conservative 46; Mismatches 84; Indels 1; Gaps 1; DR QY 5 KGLSAAEKKTRMMEIFSTKTDVFLDKLDEKIKAPKEKGITAMSVKVQSLVQSLVDDNIVTEK 66 DR 4 KGTSLDEKKERLLEMKRKGKLYIQLSLSBGSQSKHASLQSKTEAKIGRCETEERTRLAKE 124 DR 65 IGTSTYKWAFFSKALHAKHKHLVLESQSESGSKHASLQSKTEAKIGRCETEERTRLAKE 124 DR 63 IGASTYKWCASKRSQSRARTELARIQKALSRQINFIDKATARIEBKGVBETEERSSSL 122

QY	125	KELSSURDQRQKARVEVKYKCDPQVVERQANKVAKEARNRWTNDNIFAIKSWAKRKF	184	RL	Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
Db	123	KEKLALQVKLBEORGTFRDLLKNDPVAQKLUNYDIAQKANLWTNDNIFCQLQVMILK	182	DR	EMBL; AY55061; AAS9347_1; -.
Qy	185	GFEENKIDRFGIPEDDYD 205		DR	InterPro; IPR005647; Mnd1; -.
Db	183	QMDKKTIVSTALGIGTGFVDYL 203		DR	Pfam; PF03962; Mnd1; 1.
		RESULT 11		SQ	SEQUENCE 179 AA; 20900 MW; 6E00169203505476 CRC64;
Q7QX0		PRELIMINARY;	PRT; 203 AA.		
Q7QX0		PRELIMINARY;	PRT; 203 AA.		
AC	Q7QX0;	01-MAR-2004 (TREMBREL. 26, Created)			
DT	01-MAR-2004 (TREMBREL. 26, Last sequence update)				
CC	01-MAR-2004 (TREMBREL. 26, Last annotation update)				
DR	GLP-76 12374 11763				
OS	Giardia lamblia ArcC 58903.				
OC	Bukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.				
OX	NCBI_TaxID:184922;				
RN	{1}				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=WB C6;				
RA	Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,				
RA	Olsen G.J., Sogin M.L.;				
RT	"Draft sequence of the Giardia lamblia genome."				
RL	Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.				
-!- CAUTION: The sequence shown here is derived from an					
CC	EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is				
DR	preliminary data.				
EMBL; AAC000100; ERA38463; 1; -.					
DR	InterPro; IPR005647; Mnd1.				
DR	InterPro; IPR009058; Wing_hlx_DNA_bnd.				
DR	Pfam; PF03962; Mnd1; 1.				
SQ	SEQUENCE 203 AA; 23278 MW; CC0625DCDC4158C0 CRC64;				
	Query Match				
	Best Local Similarity 34.8%; Score 340.5; DB 2; Length 203;				
	Matches 70; Conservative 46; Mismatches 84; Indels 1; Gaps 1;				
Qy	5	KGLSAEKKRTRMMEIFSETKDVFOLKOLEKKIAPKEKGITAMSVKEVQSLVDDGM	59	RL	Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
Qy	6	4 :			
Q7RHS3		PRELIMINARY;	PRT; 196 AA.		
Q7RHS3		PRELIMINARY;	PRT; 196 AA.		
AC	Q7RHS3;	01-MAR-2004 (TREMBREL. 26, Created)			
CC	01-MAR-2004 (TREMBREL. 26, Last sequence update)				
DR	01-MAR-2004 (TREMBREL. 26, Last annotation update)				
DB	01-MAR-2004 (TREMBREL. 26, Last)				
DR	Homo sapiens GAJ, putative.				
DR	Name=PY0140;				
DR	Plasmodium yoelii yoelii.				
DR	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
DR	NCBI_TaxID:73239;				
DR	SEQUENCE FROM N.A.				
Q7RH53		PRELIMINARY;	PRT; 196 AA.		
Q7RH53		PRELIMINARY;	PRT; 196 AA.		
AC	Q7RH53;	Published=12368865; DOI=10.1038/nature01099;			
RA	Carlton J.M., Anguoli S.V., Subh B.B., Kooij T.W., Perreia M.,				
RA	Peterson J.D., Emiola M.D., Allen J.B., Seligut J.D., Koo H.L.,				
RA	Shallom S.J., van Aken S.B., Riedmiller S.B., Feidlyum T.V.,				
RA	Cho J.K., Quackenbush J., Sedegah M., Snaibi A., Cummings L.M.,				
RA	Florens L., Yates F.R., Ill, Raine J.D., Sinden R.E., Harris M.A.,				
RA	Cunningham D.A., Preller P.R., Bergman L.W., Vaidya A.B.,				
RA	van Lin L.H., Jame C.J., Waters A.P., Smith H.O., White O.R.,				
RA	Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,				
RA	Carucci D.J.,				
RA	RT	Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii.			
RA	RT	Nature 419:512-519 (2002).			
CC	CC	-!- CAUTION: The sequence shown here is derived from an			
DR	preliminary data.				
EMBL; AABL001237; ERA15959; 1.					
DR	InterPro; IPR005647; Mnd1.				
DR	Pfam; PF03962; Mnd1; 1.				
DR	SEQUENCE 196 AA; 23066 MW; 7FD88207329074 CRC64;				
Q7RHS4		Query Match	26.5%; Score 277; DB 2; Length 196;		
AC	Q7RHS4;	Best Local Similarity 29.7%; Pred. No. 1. 9e-09;			
DT	05-JUL-2004 (TREMBREL. 27, last sequence update)				
DT	05-JUL-2004 (TREMBREL. 27, last annotation update)				
DB	MND1 domain containing protein.				
OS	Aedes aegypti (Yellowfever mosquito).				
OC	Eukaryota; Metazoa; Arthropoda; Diptera; Nematocera; Culicoidea; Aedes.				
OC	Neoplera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.				
OC	NCBI_TaxID:7159;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=SALIVARY GLAND;				
RA	Chandra P.K., Wikel S.K.;				
QY	116	ERIGTSWYAFSKALHCKLVELEQSLI-----GSQKHSLSKIEAKIGRC	115	RL	Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
QY	63	ERIGTSWYAFSKALHCKLVELEQSLI-----GSQKHSLSKIEAKIGRC	115	DR	EMBL; AY55061; AAS9347_1; -.
QY	61	ERIGTSWYAFSKALHCKLVELEQSLI-----GSQKHSLSKIEAKIGRC	114	DR	InterPro; IPR005647; Mnd1; 1.
QY	116	ERIGTSWYAFSKALHCKLVELEQSLI-----GSQKHSLSKIEAKIGRC	114	DR	Pfam; PF03962; Mnd1; 1.

Db 115 KDELKDILKEVKNVLDKKSIEKKSELDKLKTDIQRQEMKIQSNFATESIERWNINFL 174
 Qy 176 IKSWARFKFGEEENKIDRTRGI 197
 Db 175 IJKWIDORTKNGDWDRLGM 196

RESULT 14

Q6BQJ5 PRELIMINARY; PRT; 211 AA.
 ID Q6BQJ5; PRELIMINARY; PRT; 225 AA.
 AC 075CHS; PRELIMINARY; PRT; 225 AA.
 DT 05-JUN-2004 (TREMBLrel. 27, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Similar to CA3413|IPR9239 Candida albicans IPR9239 unknown function
 DE (Fragment).
 DE ORFNames=DEHA0804999;
 OS Debaryomyces hansenii CB5767.
 OC Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
 OX NCBI_TaxID=284592;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=CB5767;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marcil C., Neuveglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchard S., Beckerich M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
 RA Desponts L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hanstrøm F., Henneguin C., Janaiaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swenene D., Tekla F., Wesołowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Boločin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts.";
 RN [2] SEQUENCE FROM N.A.
 RP STRAIN=CB5767;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR312137; CAGB7731; 1.
 DR InterPro; IPR005647; Mnd1.
 DR Pfam; PF03962; Mnd1.
 FT NON_TER 1
 SEQUENCE 211 AA; F8B5D6C618334C8 CRC64;

Query Match 21.1%; Score 221; DB 2; length 225;
 Best Local Similarity 26.4%; Pred. No. 5.3e-06; Gaps 5;
 Matches 56; Conservative 51; Mismatches 93; Indels 12; Gaps 5;

Db 3 KKKGLSEEKTRMMELFSETKDVQKDLKTIAPER-KGTTAMSTKRVQSLVD-DGMV 60
 4 KRAVVTIAEKKARVULKFOEBSITYSKDLEKIPKKCAGVSSMLVKRDIVQOLIDEGLI 63

Qy 61 DGERIGTSNYWAPSKALHARKHKLVEQLEQSEGSSQKHSLOKSE-----KAKIGR 114
 Db 64 SVEKGAVNVVTCPKQVQVGNMTCMAMKARSESSQVRLQELQAINNSEICKHARA 123

Qy 115 CTEERTRR-LAKELSSIRDREOLKAEVEVKYKDC-DPOVVEERQANKVYAKEAARNWT 170

Db 124 SFGVSYTRQALTEHDELGROLAALOSAYRKLEDITKIDSYCGRVSKUEOLDKIT 183

Qy 171 DNIAFIKSWARFKFGEEENKIDRTRGIPEDF 202
 Db 184 DNIEVIVSFLMRRAVSYRASLAALDPEEEF 215

Search completed: March 23, 2005, 19:17:52
 Job time : 180 secs

Query Match 25.5%; Score 266.5; DB 2; Length 211;
 Best Local Similarity 29.0%; Pred. No. 8.8e-09; Gaps 5;
 Matches 61; Conservative 55; Mismatches 75; Indels 19;

Qy 4 KKGLSAEKTRRMMELFSETKDVQKDLKTIAPER-KGTTAMSTKRVQSLVD-DGMV 63
 Db 3 KKGLSAEDKKEKLYEFNQSHFTYKLGKIEKKGSKYKISMLKDVQQLDNNLNC 62

Qy 64 RIGTSNYWAPSKALHARKHKLVEQLEQSEGSSQKHSLOKSE-----KAKIGR 117
 Db 63 KGTGTTNDLYWC-----KFDKIKTQDQVNNYQNKKEKQLERDQLEKIQOLGKQRLV 115

Qy 118 ---EERTRLAKLSSIRDQSKQKAEVEVKYKDCDPOVVEERQANKVYAKEAARNWT 173

Db 116 KSDFGNRHKLIDPOFTCLSKRKLHLEELLYKGNDPQLQIDERKNVHLIAETFDDI 175

Qy 174 FAKSWARF-KFGEEENKIDRTRGIPEDF 201

Db 176 ESMIVYPTKVSATIEEELDRSLBGLISEF 205

THIS PAGE BLANK (USPTO)